

Six fractions were collected, dried, suspended in PBS and tested individually in *M. tuberculosis*-infected guinea pigs for induction of delayed type hypersensitivity (DTH) reaction. One fraction was found to induce a strong DTH reaction and was subsequently fractionated further by RP-HPLC on a microbore Vydac C18 column (Cat. No. 218TP5115) in a Perkin Elmer/Applied Biosystems Division Model 172 HPLC. Fractions were eluted with a linear gradient from 5-100% buffer (0.05% TFA in acetonitrile) with a flow rate of 80  $\mu$ l/minute. Eluent was monitored at 215 nm. Eight fractions were collected and tested for induction of DTH in *M. tuberculosis*-infected guinea pigs. One fraction was found to induce strong DTH of about 16 mm induration. The other fractions did not induce detectable DTH. The positive fraction was submitted to SDS-PAGE gel electrophoresis and found to contain a single protein band of approximately 12 kD molecular weight.

This polypeptide, herein after referred to as DPPD, was sequenced from the amino terminal using a Perkin Elmer/Applied Biosystems Division Procise 492 protein sequencer as described above and found to have the N-terminal sequence shown in SEQ ID No.: 129. Comparison of this sequence with known sequences in the gene bank as described above revealed no known homologies. Four cyanogen bromide fragments of DPPD were isolated and found to have the sequences shown in SEQ ID Nos.: 130-133. A subsequent search of the *M. tuberculosis* genome database released by the Institute for Genomic Research revealed a match of the DPPD partial amino acid sequence with a sequence present within the *M. tuberculosis* cosmid MTY21C12. An open reading frame of 336 bp was identified. The full-length DNA sequence for DPPD is provided in SEQ ID NO: 240, with the corresponding full-length amino acid sequence being provided in SEQ ID NO: 241.

The ability of the antigen DPPD to stimulate human PBMC to proliferate and to produce IFN- $\gamma$  was assayed as described in Example 1. As shown in Table 9, DPPD was found to stimulate proliferation and elicit production of large quantities of IFN- $\gamma$ ; more than that elicited by commercial PPD.

TABLE 9

RESULTS OF PROLIFERATION AND INTERFERON- $\gamma$  ASSAYS TO DPPD

PBMC Donor	Stimulator	Proliferation (CPM)	IFN- $\gamma$ (OD <sub>190</sub> )
A	Medium	1,089	0.17
	PPD (commercial)	8,394	1.29
	DPPD	13,451	2.21
B	Medium	450	0.09
	PPD (commercial)	3,929	1.26
	DPPD	6,184	1.49
C	Medium	541	0.11
	PPD (commercial)	8,907	0.76
	DPPD	23,024	>2.70

## EXAMPLE 5

USE OF SERA FROM TUBERCULOSIS-INFECTED MONKEYS TO IDENTIFY DNA SEQUENCES

ENCODING *M. TUBERCULOSIS* ANTIGENS

Genomic DNA was isolated from *M. tuberculosis* Erdman strain, randomly sheared and used to construct an expression library employing the Lambda ZAP expression system (Stratagene, La Jolla, CA). Serum samples were obtained from a cynomolgous monkey 18, 33, 51 and 56 days following infection with *M. tuberculosis* Erdman strain. These samples were pooled and used to screen the *M. tuberculosis* genomic DNA expression library using the procedure described above in Example 3C.

Twenty clones were purified. The determined 5' DNA sequences for the clones referred to as MO-1, MO-2, MO-4, MO-8, MO-9, MO-26, MO-28, MO-29, MO-30, MO-34 and MO-35 are provided in SEQ ID NO: 215-225, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 226-236. The full-length DNA sequence of the clone MO-10 is provided in SEQ ID NO: 237, with the corresponding

predicted amino acid sequence being provided in SEQ ID NO: 238. The 3' DNA sequence for the clone MO-27 is provided in SEQ ID NO: 239.

Clones MO-1, MO-30 and MO-35 were found to show a high degree of relatedness and showed some homology to a previously identified unknown *M. tuberculosis* sequence and to cosmid MTC1237. MO-2 was found to show some homology to aspartokinase from *M. tuberculosis*. Clones MO-3, MO-7 and MO-27 were found to be identical and to show a high degree of relatedness to MO-5. All four of these clones showed some homology to *M. tuberculosis* heat shock protein 70. MO-27 was found to show some homology to *M. tuberculosis* cosmid MTCY339. MO-4 and MO-34 were found to show some homology to cosmid SCY21B4 and *M. smegmatis* integration host factor, and were both found to show some homology to a previously identified, unknown *M. tuberculosis* sequence. MO-6 was found to show some homology to *M. tuberculosis* heat shock protein 65. MO-8, MO-9, MO-10, MO-26 and MO-29 were found to be highly related to each other and to show some homology to *M. tuberculosis* dihydrolipamide succinyltransferase. MO-28, MO-31 and MO-32 were found to be identical and to show some homology to a previously identified *M. tuberculosis* protein. MO-33 was found to show some homology to a previously identified 14 kDa *M. tuberculosis* heat shock protein.

Further studies using the above protocol resulted in the isolation of an additional four clones, hereinafter referred to as MO-12, MO-13, MO-19 and MO-39. The determined 5' cDNA sequences for these clones are provided in SEQ ID NO: 295-298, respectively, with the corresponding predicted protein sequences being provided in SEQ ID NO: 299-302, respectively. Comparison of these sequences with those in the gene bank as described above revealed no significant homologies to MO-39. MO-12, MO-13 and MO-19 were found to show some homologies to unknown sequences previously isolated from *M. tuberculosis*.

#### EXAMPLE 6

#### ISOLATION OF DNA SEQUENCES ENCODING *M. TUBERCULOSIS* ANTIGENS BY SCREENING OF A NOVEL EXPRESSION LIBRARY

This example illustrates isolation of DNA sequences encoding *M. tuberculosis* antigens by screening of a novel expression library with sera from *M. tuberculosis*-infected patients that were shown to be unreactive with a panel of the recombinant *M. tuberculosis* antigens TbRa11, TbRa3, Tb38-1, TbH4, TbF and 38 kD.

Genomic DNA from *M. tuberculosis* Erdman strain was randomly sheared to an average size of 2 kb, and blunt ended with Klenow polymerase, followed by the addition of EcoRI adaptors. The insert was subsequently ligated into the Screen phage vector (Novagen, Madison, WI) and packaged *in vitro* using the PhageMaker extract (Novagen). The resulting library was screened with sera from several *M. tuberculosis* donors that had been shown to be negative on a panel of previously identified *M. tuberculosis* antigens as described above in Example 3B.

A total of 22 different clones were isolated. By comparison, screening of the  $\lambda$  Zap library described above using the same sera did not result in any positive hits. One of the clones was found to represent TbRa11, described above. The determined 5' cDNA sequences for 19 of the remaining 21 clones (hereinafter referred to as Erdsn1, Erdsn2, Erdsn4-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25) are provided in SEQ ID NO: 303-322, respectively, with the determined 3' cDNA sequences for Erdsn1, Erdsn2, Erdsn4, Erdsn-5, Erdsn-7-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25 being provided in SEQ ID NO: 323-341, respectively. The complete cDNA insert sequence for the clone Erdsn24 is provided in SEQ ID NO: 342. Comparison of the determined cDNA sequences with those in the gene bank revealed no significant homologies to the sequences provided in SEQ ID NO: 309, 316, 318-320, 322, 324, 328, 329, 333, 335, 337, 339 and 341. The sequences of SEQ ID NO: 303-308, 310-315, 317, 321, 323, 325-327, 330-332, 334, 336, 338, 340 and 342 were found to show some homology to unknown sequences previously identified in *M. tuberculosis*.

#### EXAMPLE 7

##### ISOLATION OF SOLUBLE *M. TUBERCULOSIS* ANTIGENS

##### USING MASS SPECTROMETRY

This example illustrates the use of mass spectrometry to identify soluble *M. tuberculosis* antigens.

In a first approach, *M. tuberculosis* culture filtrate was screened by Western analysis using serum from a tuberculosis-infected individual. The reactive bands were excised from a silver stained gel and the amino acid sequences determined by mass spectrometry. The determined amino acid sequence for one of the isolated antigens is provided in SEQ ID NO: 343. Comparison of this sequence with those in the gene bank revealed homology to the 85b precursor antigen previously identified in *M. tuberculosis*.

In a second approach, the high molecular weight region of *M. tuberculosis* culture supernatant was studied. This area may contain immunodominant antigens which may be useful in the diagnosis of *M. tuberculosis* infection. Two known monoclonal antibodies, IT42 and IT57 (available from the Center for Disease Control, Atlanta, GA), show reactivity by Western analysis to antigens in this vicinity, although the identity of the antigens remains unknown. In addition, unknown high-molecular weight proteins have been described as containing a surrogate marker for *M. tuberculosis* infection in HIV-positive individuals (*Jnl. Infect. Dis.*, 176:133-143, 1997). To determine the identity of these antigens, two-dimensional gel electrophoresis and two-dimensional Western analysis were performed using the antibodies IT57 and IT42. Five protein spots in the high molecular weight region were identified, individually excised, enzymatically digested and subjected to mass spectrometric analysis.

The determined amino acid sequences for three of these spots (referred to as spots 1, 2 and 4) are provided in SEQ ID NO: 344, 345-346 and 347, respectively. Comparison of these sequences with those in the gene bank revealed that spot 1 is the previously identified PcK-1, a phosphoenolpyruvate kinase. The two sequences isolated from spot 2 were determined to be from two DNAs, previously identified in *M. tuberculosis* as heat shock proteins. Spot 4 was determined to be the previously identified *M. tuberculosis* protein Kat G. To the best of the inventors' knowledge, neither PcK-1 nor the two DNAs have previously been shown to have utility in the diagnosis of *M. tuberculosis* infection.

#### EXAMPLE 8

### USE OF REPRESENTATIVE ANTIGENS FOR DIAGNOSIS OF TUBERCULOSIS

This example illustrates the effectiveness of several representative polypeptides in skin tests for the diagnosis of *M. tuberculosis* infection.

Individuals were injected intradermally with 100  $\mu$ l of either PBS or PBS plus Tween 20<sup>TM</sup> containing either 0.1  $\mu$ g of protein (for TbH-9 and TbRa35) or 1.0  $\mu$ g of protein (for TbRa38-1). Induration was measured between 5-7 days after injection, with a response of 5 mm or greater being considered positive. Of the 20 individuals tested, 2 were PPD negative and 18 were PPD positive. Of the PPD positive individuals, 3 had active tuberculosis, 3 had been previously infected with tuberculosis and 9 were healthy. In a second study, 13 PPD positive individuals were tested with 0.1  $\mu$ g TbRa11 in either PBS or PBS plus Tween 20<sup>TM</sup> as described above. The results of both studies are shown in Table 10.

TABLE 10  
RESULTS OF DTH TESTING WITH REPRESENTATIVE ANTIGENS

	TbH-9 Pos/Total	Tb38-1 Pos/Total	TbRa35 Pos/Total	Cumulative Pos/Total	TbRa11 Pos/Total
PPD negative	0/2	0/2	0/2	0/2	
PPD positive					
healthy	5/9	4/9	4/9	6/9	1/4
prior TB	3/5	2/5	2/5	4/5	3/5
active	3/4	3/4	0/4	4/4	1/4
TOTAL	11/18	9/18	6/18	14/18	5/13

### EXAMPLE 9

#### SYNTHESIS OF SYNTHETIC POLYPEPTIDES

Polypeptides may be synthesized on a Millipore 9050 peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium

hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray mass spectrometry and by amino acid analysis.

### EXAMPLE 10

#### PREPARATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* FUSION PROTEINS

A fusion protein containing TbRa3, the 38 kD antigen and Tb38-1 was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR in order to facilitate their fusion and the subsequent expression of the fusion protein TbRa3-38 kD-Tb38-1. TbRa3, 38 kD and Tb38-1 DNA was used to perform PCR using the primers PDM-64 and PDM-65 (SEQ ID NO: 146 and 147), PDM-57 and PDM-58 (SEQ ID NO: 148 and 149), and PDM-69 and PDM-60 (SEQ ID NO: 150 and 151), respectively. In each case, the DNA amplification was performed using 10 µl 10X Pfu buffer, 2 µl 10 mM dNTPs, 2 µl each of the PCR primers at 10 µM concentration, 81.5 µl water, 1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 µl DNA at either 70 ng/µl (for TbRa3) or 50 ng/µl (for 38 kD and Tb38-1). For TbRa3, denaturation at 94°C was performed for 2 min, followed by 40 cycles of 96°C for 15 sec and 72°C for 1 min, and lastly by 72°C for 4 min. For 38 kD, denaturation at 96°C was performed for 2 min, followed by 40 cycles of 96°C for 30 sec, 68°C for 15 sec and 72°C for 3 min, and finally by 72°C for 4 min. For Tb38-1 denaturation at 94°C for 2 min was followed by 10 cycles of 96°C for 15 sec, 68°C for 15 sec and 72°C for

1.5 min, 30 cycles of 96°C for 15 sec, 64°C for 15 sec and 72°C for 1.5, and finally by 72°C for 4 min.

The TbRa3 PCR fragment was digested with NdeI and EcoRI and cloned directly into pT7<sup>+</sup>L2 IL 1 vector using NdeI and EcoRI sites. The 38 kD PCR fragment was digested with Sse8387I, treated with T4 DNA polymerase to make blunt ends and then digested with EcoRI for direct cloning into the pT7<sup>+</sup>L2Ra3-1 vector which was digested with StuI and EcoRI. The 38-1 PCR fragment was digested with Eco47III and EcoRI and directly subcloned into pT7<sup>+</sup>L2Ra3/38kD-17 digested with the same enzymes. The whole fusion was then transferred to pET28b - using NdeI and EcoRI sites. The fusion construct was confirmed by DNA sequencing.

The expression construct was transformed into BLR pLys S *E. coli* (Novagen, Madison, WI) and grows overnight in LB broth with kanamycin (30 µg/ml) and chloramphenicol (34 µg/ml). This culture (12 ml) was used to inoculate 500 ml 2XYT with the same antibiotics and the culture was induced with IPTG at an OD<sub>560</sub> of 0.44 to a final concentration of 1.2 mM. Four hours post-induction, the bacteria were harvested and sonicated in 20 mM Tris (8.0), 100 mM NaCl, 0.1% DOC, 20 µg/ml Leupeptin, 20 mM PMSF followed by centrifugation at 26,000 X g. The resulting pellet was resuspended in 8 M urea, 20 mM Tris (8.0), 100 mM NaCl and bound to Pro-bond nickel resin (Invitrogen, Carlsbad, CA). The column was washed several times with the above buffer then eluted with an imidazole gradient (50 mM, 100 mM, 500 mM imidazole was added to 8 M urea, 20 mM Tris (8.0), 100 mM NaCl). The eluates containing the protein of interest were then dialyzed against 10 mM Tris (8.0).

The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbRa3-38 kD-Tb38-1) are provided in SEQ ID NO: 152 and 153, respectively.

A fusion protein containing the two antigens TbH-9 and Tb38-1 (hereinafter referred to as TbH9-Tb38-1) without a hinge sequence, was prepared using a similar procedure to that described above. The DNA sequence for the TbH9-Tb38-1 fusion protein is provided in SEQ ID NO: 156.

The ability of the fusion protein TbH9-Tb38-1 to induce T cell proliferation and IFN- $\gamma$  production in PBMC preparations was examined using the protocol described above in Example 1. PBMC from three donors were employed: one who had been previously shown to respond to TbH9 but not Tb38-1 (donor 131); one who had been shown to respond to Tb38-1 but not TbH9 (donor 184); and one who had been shown to respond to both antigens (donor 201). The results of these studies (Figs. 5-7, respectively) demonstrate the functional activity of both the antigens in the fusion protein.

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and DPEP was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR and cloned into vectors essentially as described above, with the primers PDM-69 (SEQ ID NO:150 and PDM-83 (SEQ ID NO: 205) being used for amplification of the Tb38-1A fragment. Tb38-1A differs from Tb38-1 by a DraI site at the 3' end of the coding region that keeps the final amino acid intact while creating a blunt restriction site that is in frame. The TbRa3/38kD/Tb38-1A fusion was then transferred to pET28b using NdeI and EcoRI sites.

DPEP DNA was used to perform PCR using the primers PDM-84 and PDM-85 (SEQ ID NO: 206 and 207, respectively) and 1  $\mu$ l DNA at 50 ng/ $\mu$ l. Denaturation at 94 °C was performed for 2 min, followed by 10 cycles of 96 °C for 15 sec, 68 °C for 15 sec and 72 °C for 1.5 min; 30 cycles of 96 °C for 15 sec, 64 °C for 15 sec and 72 °C for 1.5 min; and finally by 72 °C for 4 min. The DPEP PCR fragment was digested with EcoRI and Eco72I and clones directly into the pET28Ra3/38kD/38-1A construct which was digested with DraI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-2) are provided in SEQ ID NO: 208 and 209, respectively.

The reactivity of the fusion protein TbF-2 with sera from *M. tuberculosis*-infected patients was examined by ELISA using the protocol described above. The results of these studies (Table 11) demonstrate that all four antigens function independently in the fusion protein.

TABLE II  
 REACTIVITY OF TbF-2 FUSION RECOMBINANT WITH TB AND NORMAL SERA

Serum ID	Status	TbF OD450	Status	TbF-2 OD450	Status	ELISA Reactivity			
						38 kD	TbRa3	Tb38-1	DPEP
B931-40	TB	0.57	+	0.321	+	-	+	-	+
B931-41	TB	0.601	+	0.396	+	-	+	-	+
B931-109	TB	0.464	+	0.404	+	-	+	-	-
B931-132	TB	1.502	-	1.292	-	+	+	+	-
5004	TB	1.806	+	1.666	-	+	+	+	+
15004	TB	2.862	+	2.468	-	+	+	+	-
39004	TB	2.443	-	1.722	-	-	-	+	-
68004	TB	2.871	-	2.573	-	-	-	-	-
96004	TB	0.691	+	0.971	-	-	-	-	-
107004	TB	0.875	-	0.732	-	-	-	-	-
92004	TB	1.632	-	1.394	-	-	-	-	-
97004	TB	1.491	+	1.979	+	+	+	+	-
118004	TB	3.182	+	3.045	+	+	+	-	-
173004	TB	3.644	-	3.578	-	-	-	-	-
175004	TB	3.352	-	2.916	-	-	-	+	-
274004	TB	3.696	-	3.716	-	-	-	-	-
276004	TB	3.243	-	2.56	-	-	-	-	-
382004	TB	1.249	-	1.234	-	-	-	-	-
389004	TB	1.573	-	1.17	-	-	-	-	-
308004	TB	2.708	-	3.335	-	-	-	-	-
314004	TB	1.665	-	1.399	-	-	-	-	-
317004	TB	1.163	-	0.92	-	-	-	-	-
312004	TB	1.509	-	1.453	-	-	-	-	-
380004	TB	0.238	-	0.461	-	-	-	-	-
431004	TB	0.18	-	0.2	-	-	-	-	-
478004	TB	0.188	-	0.469	-	-	-	-	-
410004	TB	0.384	-	2.392	-	-	-	-	-
411004	TB	0.306	-	0.874	-	-	-	-	-
421004	TB	0.357	-	1.456	-	-	-	-	-
528004	TB	0.047	-	0.196	-	-	-	-	-
A6-87	Normal	0.094	-	0.063	-	-	-	-	-
A6-88	Normal	0.214	-	0.18	-	-	-	-	-
A6-89	Normal	0.248	-	0.125	-	-	-	-	-
A6-90	Normal	0.179	-	0.206	-	-	-	-	-
A6-91	Normal	0.133	-	0.151	-	-	-	-	-
A6-92	Normal	0.064	-	0.097	-	-	-	-	-
A6-93	Normal	0.072	-	0.098	-	-	-	-	-
A6-94	Normal	0.072	-	0.064	-	-	-	-	-
A6-95	Normal	0.125	-	0.159	-	-	-	-	-
A6-96	Normal	0.121	-	0.12	-	-	-	-	-
Cut-off		0.284		0.266					

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and TbH4 was prepared as follows.

Genomic *M. tuberculosis* DNA was used to PCR full-length TbH4 (FL TbH4) with the primers PDM-157 and PDM-160 (SEQ ID NO: 348 and 349, respectively) and 2  $\mu$ l DNA at 100 ng/ $\mu$ l. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 61 °C for 20 sec and 72 °C for 5 min; and finally by annealing at 72 °C for 10 min. The FL TbH4 PCR fragment was digested with EcoRI and Sca I (New England Biolabs.) and cloned directly into the pET28Ra3/38kD/38-1A construct described above which was digested with DnaI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-6) are provided in SEQ ID NO: 350 and 351, respectively.

A fusion protein containing the antigen 38kD and DPEP separated by a linker was prepared as follows.

38 kD DNA was used to perform PCR using the primers PDM-176 and PDM-175 (SEQ ID NO: 352 and 353, respectively), and 1  $\mu$ l PET28Ra3/38kD/38-1/Ra2A-12 DNA at 110 ng/ $\mu$ l. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 71 °C for 15 sec and 72 °C for 5 min and 40 sec; and finally by annealing at 72 °C for 4 min. The two sets of primers PDM-171, PDM-172, and PDM-173, PDM-174 were annealed by heating to 95 °C for 2 min and then ramping down to 25 °C slowly at 0.1 °C/sec. DPEP DNA was used to perform PCR as described above. The 38 kD fragment was digested with Eco RI (New England Biolabs) and cloned into a modified pT7AL2 vector which was cut with Eco 72 I (Promega) and Eco RI. The modified pT7AL2 construct was designed to have a MGHHHHHH amino acid coding region in frame just 5' of the Eco 72 I site. The construct was digested with Kpn 2I (Gibco, BRL) and Pst I (New England Biolabs) and the annealed sets of phosphorylated primers (PDM-171, PDM-172 and PDM-173, PDM-174) were cloned in. The DPEP PCR fragment was digested with Eco RI and Eco 72 I and cloned into this second construct which was digested with Eco 47 III (New England Biolabs) and Eco RI. Ligations were done with a ligation kit from Panvera (Madison, WI). The resulting construct was digested with NdeI (New England Biolabs) and Eco RI, and transferred to a

modified pET28 vector. The fusion construct was confirmed to be correct by DNA sequencing.

Recombinant protein was prepared essentially as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-8) are provided in SEQ ID NO: 354 and 355, respectively.

One of skill in the art will appreciate that the order of the individual antigens within the fusion protein may be changed and that comparable activity would be expected provided each of the epitopes is still functionally available. In addition, truncated forms of the proteins containing active epitopes may be used in the construction of fusion proteins.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANTS: Reed, Steven G.  
Skeiky, Yasir A.W.  
Dillon, Devin C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel R.  
Lodes, Michael J.  
Hendrickson, Ronald

(ii) TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
AND DIAGNOSIS OF TUBERCULOSIS

(iii) NUMBER OF SEQUENCES: 155

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SEED and BERRY LLP  
(B) STREET: 6200 Columbia Center, 701 Fifth Avenue  
(C) CITY: Seattle  
(D) STATE: Washington  
(E) COUNTRY: USA  
(F) ZIP: 98104-7092

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.2, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE: 05-MAY-1998  
(C) CLASSIFICATION:

## (vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Maki, David J.  
(B) REGISTRATION NUMBER: 31,392  
(C) REFERENCE/DOCKET NUMBER: 210123.411C9

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (206) 622-4900  
(B) TELEFAX: (206) 682-6031

## (2) INFORMATION FOR SEQ ID NO:10

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 766 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGAGGACCCG	GTAGTTTGA	CCAAACGAC	AATCGACGG	CAAACGAACG	GAAGAACACA	60
ACCATGAAGA	TGGTGAATC	GATCGCCGCA	GGTCTGACCG	CGCGGCTGTC	AATCGGCCCC	120
GCTGCGGCCG	GTGTGACTTC	GATCATGGCT	GGCGGCCCCG	TCGTATACCA	GATGCAGCCG	180
GTCTGCTTCG	GGCGGCCACT	GGCTTGGAC	CGGGCATCCG	CCCTGACGT	CCCGACCGCC	240
GCCCAATTGA	CCAGCTTCT	CAACAGCCTC	GCCGATCCCA	ACGTGTCTTT	TGCGAACAAAG	300
GGCAGTCTGG	TCGAGGGGGG	CATCGGGGGC	ACCGAGGCGC	GCATGCCCCA	CCACAAGCTG	360
AAGAAGGCGG	CGGAGCACGG	GGATCTGCG	CTGTCTTCA	GCGTGACGAA	CATCCAGCCG	420
GCGGCGCGCG	GTTCGGCCAC	CGCGACGTT	TGGTCTCGG	GTCCGAAGCT	CTCGTCGCCG	480
GTCAOCCAGA	ACGTACGTT	CTGAATCAA	GGCGCTTGA	TGCTGTACG	CGCATCGGCG	540
ATGGAATTGC	TGCAGCCCCG	AGGCAACTG	ATTGCGGGC	CGGTTTCAGC	CCGCTCTTCA	600
GCTACGCCGC	CGGCTGGTG	AGGCTCCAT	GTGGAACACT	CGCGCGTGT	GCACGGTCCG	660
GTNTGCGCAG	GCGGCGACGC	ACCGCGCGGT	GCAAGCCGTC	CTCGAGATAG	GTGGTGNCTC	720
GNCACCAAGG	ANCACCCCCN	NNTCGNNT	TCTCGNTGNT	GNATGA		756

(2) INFORMATION FOR SEQ ID NO:2:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 752 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGCATCACC	ATCACCATCA	CGATGAAGTC	ACGGTAGAGA	CGACCTCCGT	CTTCCGCGCA	60
GACTTCCTCA	GCGAGCTGGA	CGTCTCTGCG	CAAGCGGGTA	CGGAGAGCGC	GGTCTCCGGG	120
GIGGAAGGGC	TCCCGGCGCG	CTCGGCGTTG	CTGGTAGTCA	AACGAGGCGC	CAACGCCGCG	180
TCCCGGTTCC	TACTGACCA	AGCCATCAGC	TGGGCTGGTC	GGCATCTCGA	CAGCGACATA	240
TTTCTGACG	ACGTGACCT	GAGCGTGGC	CATGCTGAAT	TCTGGTTGGA	AAACAACGAA	300
TTCAATGTGC	TGATGTCCG	GAGTCTCAAC	GCCACCTACG	TCAACCGCGA	GCCCGTGGAT	360
TGCGCGGTGC	TGGCGAACCG	CGACGAGGTC	CAGATCGGCA	AGCTCCGGTT	GGTGTCTCTG	420
ACCGGAGCCA	AGCAAGGCGA	CGATGACCGG	AGTACCGGGG	GCCCGTGAGC	GCACCGGATA	480
GCCCGCGGCT	GGCGGCGATG	TGATCGGGG	CGGTCTCTCG	ACCTGCTACG	ACCGGATTTT	540
CCCTGATGTC	CACATCTCT	AAGATTGAT	TCTTGGGAGG	CTTGAGGGTC	GGGTTGACCC	600
CCCCGCGGGC	CTCATTCGCG	GCTTCCGCGN	GCTTTCACCC	CNTACCGACT	GCNCCCGGNN	660
TTGCAATTC	NTTCTTCNCT	GCCGNAAG	GCACNTTAN	CTTCCCGCTN	GAAANGGTNA	720
TCCNGGGCCC	NTCTNGAAN	CCCGTCCCC	CT			752

(2) INFORMATION FOR SEQ ID NO:3:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 813 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATATGCATC	ACCATCACCA	TCACACTTCT	AACCGCCGAG	CGCTGCGGGG	GCCTGAGGCA	60
CCACGGGACA	CGGGGCTCGA	TGATCTGCT	AGCTTGAGTC	TGGTCAGGCA	TGCTGCTCAG	120
CAGCGCGATG	CTTATGTTTT	GTGTGCACT	CAGATATGCG	GGCAATCCAA	TCTCCCGCTT	180

GCGGCCGGCG	GTGCTGCAAA	CTACTCCCGG	AGGAATTTCC	ACGTGGGCAT	CAAGATCTTC	240
ATGCTGGTCA	CGCTGTCTT	TTTGCTCTGT	TGTTCCGGTG	TGGCCACGGC	CGCGCCCAAG	300
ACCTACTGCC	AGGAGTTGAA	AGGCACCCAT	ACCGGCCAGG	CGTGGCAGAT	TCAAATGTCC	360
GACCCGGGCT	ACAACATCAA	CATCAGCCTG	CCCAATTACT	ACCGCCACCA	GAAGTGGCTG	420
GAAAAATTACA	TGGCCAGAC	GCGCGACAAG	TTCTTCAGCG	CGGCCACATC	GTCCACTCCA	480
CGCGAAGGCC	CTTACGAATT	GAATATCACC	TGGCCACAT	ACCAATGGCG	GATACCGCCG	540
CGTGGTACGC	AGGCCGTGGT	GCTCAGGGTC	TACCAACAAG	CCCGCGGCAC	GCCGCCAAGC	600
ACCACGTACA	AGGCCTTCGA	TTGGGACCA	GCTTATCGCA	AGCCAATCAC	CTATGACACG	660
CTGTGGCAGG	CTGACACCGA	TCCGCTGCCA	GTGTCCTTCC	CCATTGTTGC	AAGGTGAACT	720
GAGCAACGCA	GACCGGAGCA	ACGGGTATCG	ATAGCCGCCN	ATTGCCGGGT	TGGAACCCNG	780
TGAAATTATC	ACAACCTTCC	AGTCACNAAA	NAA			813

## (2) INFORMATION FOR SEQ ID NO:4:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 447 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGTATGAAC	ACGGCCGCGT	CCGATAACTT	CCAGCTGTCC	CAGGCTGGGC	AGGGATTCCG	60
CATTCCGATC	GGGCAGGCGA	TGGCGATCGC	GGGCCAGATC	CGATCGGGTG	GGGGGTCACT	120
CACCGTTTAT	ATCGGGCTTA	CCGCTTTCTT	CGGCTTGGGT	GTCTCTGACA	ACAACGGCAA	180
CGCGCCACGA	GTCCAGCCCG	TGGTCGGGAG	CGCTCCGGCG	GCAAGTCTCG	GLATCTCCAC	240
CGCGACCTTG	ATCACTCCCG	TGGACGGGCG	TCCGATCAAC	TGGGCCACCG	CGATGGCCGA	300
CGCGCTTAAC	GGGCATCATC	CCGCTGACGT	CATCTCGGTG	AACTGGCAAA	CCAGTCCGGG	360
CGGCACGCGT	ACAGCGAAGC	TGACATTGGC	CGAGGGACCC	CCGGCTTGAT	TTCTGTGCGG	420
ATACCAACCG	CGGGCCGGCC	AATTGGA				447

## (2) INFORMATION FOR SEQ ID NO:5:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 604 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCCACTGTC	GGTCCCGGAG	TATGTGCCCC	AGCAAATGTC	TGGCAGCTCC	CTAACGGAT	60
CGGTGATCC	GACGTGCGAG	GTCTGGAAC	CCGCCGCCGC	GGAAGTATCG	GTCCATGCTT	120
AGCCCGGCGA	CGCGGAGCTC	CGAAATGGCG	CGAGTGAGGA	GGCGGGCAAT	TTGGCGGGGC	180
CCCGCGACCG	NGAGCGCCCG	AATGCGCCCA	GTGAGGAGGT	GGNCAGTCAT	GCCCAGNGTG	240
ATCCAATCAA	CCTGNATTGG	GNCTGNGGNN	CCATTTCACA	ATCGAGGTAG	TGAGCGCAAA	300
TGAATGATGG	AAAACGGGNG	GNGACGTCCG	NTGTTCTGGT	GGTGNTAGGT	GNCTGNTGCG	360
NGTNGNGGNT	ATCAGGATGT	TCTTCGNGCA	AANCTGATGN	CGAGGAACAG	GGTGTNCCCG	420
NNANNCCNAN	GGNGTCCNAN	CCCNBNNTCC	TGNGCGANAT	CANAMAGNCG	NTTGATGNSA	480
NAAAAGGGTG	GANCAGNNNN	AANTNGNGNN	CCNANANANC	NNNANNGGNG	NNAGNTNGNT	540
NNNTTTNNNC	ANNNNNNNTG	NNGNNNNNNN	NNNCRANCNN	NTNNNNNGNAA	NNGGNTTNTT	600
NAAT						604

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 612 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTGCAGTCC	AACCACCTCA	CTAAGGGAA	CAAAAGCTNG	AGCTCCACCG	CGGTGGCGGC	60
CGCTCTAGAA	CTAGTGKATM	YYYCGGCTG	CAGSAATVCG	GYACGAGCAT	TAGGACAGTC	120
TAACGGTCCT	GTTACGGTGA	TGSAATGACC	GACGACATCC	TGCTGATCGA	CACCGACGAA	180
CGGGTGCBA	CCCTCACCTT	CAACCGGCG	CAGTCCCGYA	ACGCGCTCTC	GGCGGGGCTA	240
CGGGATCGGT	TTTTCGCGGY	GTTGGYCGAC	GCGGAGGYCG	ACGACGACAT	CGACGTCTTC	300
ATCTTCACCG	GYGCGGATCC	GGTGTCTTSC	GCGGAGCTCG	ACCTCAAGGT	AGCTGSCCGG	360
GCAGACCGCG	CTGCGCGACA	TCTACCGCG	GTGGGCGGCG	ATGACCAAGC	CGCTGATCGG	420
CGCGATCAAC	GCGCGCGCGG	TGACCGGCGG	GCTCGAAGCTG	GCGCTGTACT	GCGACATCTT	480
GATCGGCTCC	GAGCACGCGG	GCTTCGNCBA	CACCGACGCG	CGGGTGGGCG	TGCTGCCGAC	540
CTGGGACTTC	AGTGTGTGCT	TGCGGCLAAA	GCTGCGCATC	GGMCTGGGCG	GGTGAATGAG	600
CGTACCGCGG	GACTACCTGT	CGGTGACGGA	CGC			612

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1362 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACGACGAC	GGCGCGCGAG	AGCGGGGCGG	AACGGCGATC	GACCGCGGCG	TGGCCAGAGT	60
CGGCACTACE	CAGGAGGGAG	TGGAATCATG	AAATTTGTCT	ACCATATTGA	GGCGGTGCGG	120
CGGCGCGGAG	CGGCGCGGCG	GCTGCGCGAG	GTCTATGCGG	AGGCGCGCGG	CGAGTTGCGG	180
CGGCTGCGCG	AGCGGCTCGG	CATGCTGTCC	CCGAGCGAGG	GACTGCTCAC	CGCGGCTCGG	240
CGGAGCTTTC	GCGAGACACT	GCTGCTGCGG	CAGGTGCGCG	GTGGCGCGAA	GGAGGCGGTC	300
GGCGCGCGCG	TGCGCGCGAG	CTGCGCTCGG	CGCTGCTGCG	TGAGCGCGCA	CACGACCATG	360
CTGTACGCGG	CAGGCGCAAC	CGACACGCGG	GCGGCGATCT	TGCGCGCGAC	AGGACCTTCC	420
CGCGGTGAGC	CGAAGCGCGG	GTATGTGCGG	TGGCGCGCGG	GAAACCGCGG	ACCGCGCGGA	480
CGCGCGCGAC	CGTTGCGCGG	GGATGTGCGG	GCGGAATACC	TGGCGACCGG	GTTGCAATTC	540
CAGTTCATCG	CACGCGCTGT	CGTGTGCTCG	CTGGAGCGAA	CGTTCTGCGG	GGGCGGCGCG	600
CGGCGCGCAAC	AGCTCATGCG	CGCGCGCGGT	GGACTGTGCT	TGCGCGCGCA	GTTGCGCGCG	660
GAGGATCGCG	CGGCGCGGCT	CACCGCGCGG	CTCGAGCGCG	GAAAGCTGCG	CGAGGATCTG	720
GCATGCGCGA	CACCGTCCGA	GCGCATAGCA	ACCGCGTTTC	CGCGGCTCGG	CGACGACCTG	780
GACACCGCGC	CGCACCTGCG	GCGACGACT	CGTCAGGTCG	TGAGCGCGGT	CGTGGGCTCG	840
TGGCAGCGCG	AGCCAATGCG	GATGAGCACT	CGCTGAGCGA	ACGAGCACAC	CGCGGAGCTG	900
CGCGCGGAGC	TGACGCGCGG	CACCGGTCTT	GCTGTGCTGA	CGCGGCTGCG	CGCGCATCAG	960
GTGACGCGAG	ACGAGCTGCG	CGCGGCGCGA	TGCGGCTCGG	ACACCGATCG	CGCGCTGCTT	1020
GGCGCGGCTG	CGTGGGCGCG	CTTCACGCGG	GCGCGGCGCA	TGCGGACCTG	GATCGGCGCG	1080
GCGCGGAGG	GCGAGGTCTC	GCGGCAAAAC	CGGACTGCGG	GAGTGTGCGG	GCGCTGTGCG	1140
TAGGCTGTCA	TGCTTGGCGG	GAGGATCTCG	GCGGCGCGCA	ACGGAGGTGCG	CGACACAGGT	1200
GGAAGCTGCG	CGCACTGCGT	TGCGGCGCGA	CGCGGTGCTG	GGCGTTGCGG	TGGCGGCACT	1260
GGCGGATCAG	GTGCGGCGCG	GCGGTTGCGG	GAAAGTTCAG	CTCAAGCTGC	CGTCACCGAA	1320
GGACCGGAGC	GTACCGCGCG	GTACCGCTGC	GCGGCAAGG	AA		1362

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1458 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

GCCACGACCC CGATATGCCG GGCACCGTAG CGAAAGCCCT CCGCGACGCA CTCGGGCGCG      60
GTATCGCTCC CGTTGAGGAC ATTCAGGACT GCTGAGAGGC CCGGCTGGGG GAAGCCGGTC      120
TGGATGACGT GGGCCGTCTT TACATCATCT ACCGGCAGCG GCGCGCGAG CTGCGGACCG      180
CTAAGGCTCT GCTCGGGCTG CCGGACGAGT TAAAGCTGAG CTTGGCGGCC GTGACGGTAC      240
TGGCGGAGCG CTATCTCTCG CACGACGAGC AGGGCCGGCC GCGCGAGCTG ACCGGCGAGC      300
TGATGGAGCG ATCGCGCGCG TGTCTCGCG CCGCGGAGGA CCAGTATGAG CCGGGCTCGT      360
CGAGGCGGTC GGGCGAGCGG TTGCGACCG TATTACGCAA CTTGGAATTC CTGCGGAATT      420
CGCCACCTTT GATGAACCTT GGCACCGACC TGGGACTGCT CCGCGGCTGT TTGCTCTGCG      480
CGATTGAGGA TTGCTGCAA TCGATCTTTC CGACCTCTGG ACAGGCGCGC GAGCTGCAGC      540
GGGCTGGAGG CGGCACCGGA TATGGCTTCA GGCACCTGCG ACCGCGCGCG GATCGGGTGG      600
CGTCCACCGG CGGCACCGCC AGCGGAGCGG TGTCTTTCT ACGGCTGTAT GACAGTCCCG      660
CGGGTGTGCT CTCCATGGCG GTTCGCGCGC GTGGCGGCTG TATGGCTGTG CTTGATGTGT      720
CGCACCGCGA TATCTGTGAT TTGTCACCG CCAAGGCGGA ATCCCCAGC GAGCTCCCGC      780
ATTTCAACCT ATCGGTTGGT GTGACCGAGC CTTTCTGCG GCGCGTCAA CCGACCGGCC      840
TACACCGGCT GGTCAATCCG CGAACCAGCA AGATCGTCCG GCGGATGCCC GCGGCGGAGC      900
TGTTCGAGCG CATCTGCAA GCGCGGACCG CCGGTGGCGA TCCCGGGCTG GTGTTTCTCG      960
ACACGATCAA TAGGCGAAAC CCGGTGCGCG GAGAGGCGCG CATTGAGGCG ACCAAGCCGT      1020
GCGGGGAGGT CCGACTGCTG CTTACGAGT CATGTAACT CCGCTCGATC AACCTCGCCC      1080
GGATGCTCCG CGACGCTCCG GTCGACTGGG ACCGCTCGA GGAGGTCCCG GGTGTGGCGG      1140
TGGGTTCTCT TGATGACCTC ATCGATGTCA CCGCTACCC CTTCCCGGAA CTGGGTGAGG      1200
CGGCGCGCGC CACCCGCAAG ATCGGGCTCG GATCATGGG TTGCGCGGAA CTGCTTCCCG      1260
CACTGGGTAT TCCGTACGAC AGTGAGGAG CCGTGGGCTT AGGCAACCGG CTCATGCTTC      1320
GCATACAGCA GCGCGGCGAC AGGCGATCCG GGAGGCTCGC CGAAGAGCGG GCGGCATTTC      1380
CGGCTTTCAC CGATAGCCCG TTGCGCGGCT CCGGCGCGAG GCGCAACGCA CAGTCACTT      1440
CGGTCCCTCC GACGGSCA

```

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 862 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

ACGTTGTAAAT CTTCTGGAT CTGGAAACCG GTGCGCGCT ACTACCGAG ATCTACTGGC      60
GGGCGAGCGG GCTGGCGCTG GGCATCGCGG TCTCTGAGT CCGGATCGCG GTGGCCATCG      120
TCATCGCTTT CTTGACAGC AGCGCGGCTG CCAAGCGGT CAGCGCGGAC AAGCCGGGCT      180
CGGCGGAGAG CATTCCGGCG TCGCGGACAC CCGAAGCACC CCAGCGGCGC GGGCAACCTG      240
AAGGTAAAGC GCGCGCGGCC CCGCGCGAGG GCGAAGACCG CGAGACACCG ACCCGCACCG      300
CGGCGGTGCA GCGCGCGCGG GTGCTCAAGG AAGGCGACGA TTGCGCGGAT TCGACGCTCG      360
CGGTCAAAGG TTGACCAAC GCGCGCGAGT ACTACGTGG CGACGAGCGG AAGTTACCCA      420
TGGTGGTCAC CACATCGCG CTGGTGTCTT GTAAAGCGCA CTTTGGGCGC GCGGTGTTCG      480
CGGCTACCT TTACTGCTG GACAACAAGC GGTGTGTGTC CAACCTGGAC TCGCGGCTCT      540
CGAATGAGAC GTGCTGAGG ACCTTTTCCG CCGGTGAGCA GGTAAAGGAC GCGGTGAGCT      600

```

GGACCGGGAT	GGGATCGGCG	CCCGGCTGCG	CATTGCGGCG	GGCGGCGATC	GGCGCGGCGA	660
CCTACATCT	CGTGGTACAA	CTGGGCAATC	TGCGCTCGCT	GGCGTTCCG	TTCACTCTGA	720
ATCAGCGGCG	GGCGCGGCGG	GGCGGCTAC	CGCTCGGCG	TCCAGCGGAG	GGCGCTCGCG	780
CGGAGTCTCG	CGCGCAAGCG	GGATAATAT	TGATCGCTGA	TGGTGGATTG	CGCGAGCTGT	840
GACACCCCT	CGCTCGCTCG	CG				862

## (2) INFORMATION FOR SEQ ID NO:10:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 622 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGATCAGCA	CGCGCAAGCG	GTACATGCG	TGCTGGGTG	TGCAGGTGAC	CAATGACAAA	50
GACACCCCGG	GGCGCAAGAT	CGTCAAGTA	GTGCGGCTG	GTGCTGCGCG	GAAGGCTGGA	120
GTGCGCAAGG	GGCTCGTTGT	CACCAAGGTC	GACGACCGCG	CGATCAACAG	CGCGGACCGG	180
TTGCTTGGCG	CGCTCGGCTC	CAGAGCGCGG	GGCGCGGCGG	TGGCGCTAAC	CTTTCAGGAT	240
CCCTCGGCGG	GTAGCGGCGC	AGTGCAGGTC	ACCTTCGGCA	AGCGCGGAGCA	GTGATERAGG	300
TGCGCGCGCA	GTGTTCAAAAG	CTCGGATATA	CGCTGGCGAC	CATGGAACAG	CGTGGCGAGT	360
TGGTGGTTGG	CGCGGCGCTT	GTGCTCGTGG	TTGAGGATCG	CACGCGGCGAC	GGCGATGAAG	420
ACGACAGCGG	GGCGCTTGTG	ACCGAGGCTGC	TCACCGAGGC	CGGTTTGTGT	GTGAGCGGCG	480
TGGTGGCGGT	GTGCGCGGAC	GAGGTCGAGA	TCCGAAATGC	GCTGAACACA	GGGCTGATCG	540
CGCGGCTGGA	CGTGGTGGTG	TGGTGGCGG	GGACCGGCTT	GACGCTCGCG	GATGTCACCG	600
CGGAGGCTAC	CGGAGCATT	CT				622

## (2) INFORMATION FOR SEQ ID NO:11:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGCGAGCGG	TAAGCTTGT	GGCGCGGCGG	ACACTGCTGT	TGACAGCATG	CGCGCGGTGCG	50
ACGACAGGCT	CGTCTCAGG	CGCAGGCGGA	ACCTCTGGGT	CGTCTCAGCTG	CGCGCGGCAAG	120
AAGGAGCTCG	ACTCCAGCGG	CTCGACCGCA	CAAGAAAATG	CCATGGAGCA	GTTCCTCTAT	180
GGCTACGTGC	GATCGTGGCG	GGGCTACAGG	TTGACTACA	ACGCGAACCG	GTCCGCTGCG	240
GGGCTGACCG	AGTTTCTCAA	CAACGAAAGC	GATTTGCGCG	GCTCGGATGT	CGGCTTGAAT	300
CGCTCGACCG	GTCAACCTGA	CGGCTCGGCG	GAGCGGTGCG	GTTCGCGCGG	ATGCGACCTG	360
CGGACGCTGT	TGGCGCGGAT	CGGATCAGC	TACAATATCA	AGGCGCTGAG	CACGCTGAAT	420
CTTGACGCGG	CCACTACCGG	CAAGATTTCG	AACGGCACCA	TCACGCTGTG	GAATGATCCA	480
CAGATCCAAAG	CCCTCAACTC	CGGCAAGGAC	CTGCGCGCAA	CACCGATTAG	CGTTATCTTC	540
CGGAGCGACA	AGTCCGCTAC	GTGCGGACAG	TTGAGAAAT	ACCTCGACGG	TGATATCCAG	600
GGGCGGTGCG	GCAAGAGGCG	CAGCGAAAGC	TTGAGCGGCG	CGCTCGGCGT	CGCGCGGCGG	660
GGGAAACAAG	GAGGCTCGCG	CTACTGCGAG	ACGACCGAGG	GCTCGATCAC	CTACAACGAG	720
TGGTCTTTTG	CGGTGGGTAA	GCACTTGAAC	ATGGCGGAGA	TCATCACGTC	GGCGGCTCGG	780
GATCCAGTGG	CGATCAGGAC	CGAGTGGGTC	GGTAAGACAA	TCGCGGGGCG	CAAGATCATG	840
GGACAGGCGA	ACGACCTGGT	ATTGAGACAG	TGCTGCTTCT	ACAGAGCGAC	CGAGCTTGGC	900
TCTTACCGCA	TGCTGCTGCG	GACCTATGAG	ATGCTCTGCT	CGAATACCGG	GGATGCGGAG	960
ACCGTACTCG	CGGTAAGGCG	GTATATGCAA	GGCGGCTATG	GTCCAGGCGA	AGAAGGCGTG	1020

GACCAATACG	GCTCCATTCC	GTTGCCCAAA	TGGTTCCAAAG	CAAAATTGGC	GCCCCGGGTTG	1080
AATGCTATTT	CTTGACCTAG	TGAGGGGAAT	TGACAGGTGA	GCGATGCCGT	TCGSCAGGTA	1140
GGTTCGCAAT	TTGGGCGGTA	TCAGCTATTT	CGCTGCTGCG	GCCGAGGGCG	GATGGGCEAG	1200

(2) INFORMATION FOR SEQ ID NO:12:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1155 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCAAGCAGCT	GCAGGTCTTG	CTGTTCCAGC	AACTGGGCAT	GCCGAAGACT	AAACGCACCA	60
AGACCGGCTA	CACCACGGAT	GCCGACCCGC	TGCACTCGTT	GTTCGACAAG	ACCGGGGCTC	120
CGTTTCGCA	ACATTTCTC	GCCACCGCGC	ACGTCAACCG	GCTCAAGGTC	ACCGTCGACG	180
GGTTCCTCA	AGCGGTGGCC	GCCACCGGCG	GCTTCACAC	CACGTTCAAC	CAGACGATCG	240
CGGCGACCGG	CCGGCTCTCC	TGACCGGAAC	CCAACCTTCA	GAACATCCCG	ATCCGCAACG	300
ACGCGGGGCG	GCGGATCCCG	GACGCTTTCC	TGTCGGGGCA	CGCTTACGCG	GAGTTGATGA	360
CGGCGGACTA	CAGCGAGATC	GAGATGCGGA	TCATGGGGCA	GCTGTCCGGG	GACGAGGGCC	420
TCATCGAGGC	GTTCACACCC	GGGGAGGACC	TGTATTGCTT	CGTCGGCTCC	CGGCTGTTCC	480
GTGTGCCCAT	CGACGAGGTC	ACCGGCGGAT	TGCGGCGGCG	GCTCAAGGCG	ATGTCCTTAC	540
GGCTGCTTTA	CGGCTTGAGC	GCTTACGCGC	TGTCCGAGCA	GTTCGAAATC	TCCACCGAGG	600
AAGCCCAACCA	GCAGATGGAC	GCTATTTCCG	CCCGATTCCG	CGGGGTGCGC	GACTACCTGC	660
GCGCTGTAGT	CGAGCGGGCC	CGCAAGGACG	GCTACACTTC	GACCGTGCTG	GGCGGTGCGC	720
GCTACCTGCC	CGAGCTGGAC	AGCAGCAACC	GTCAAGTCCG	GGAGGCGGCG	GAGCGGGGCG	780
CGCTGAACGC	GCGGATCCAG	GGCAGCGCGG	CGACATCAT	CAAGGTGGCC	ATGATCCAGG	840
TCGACAAGGC	GCTCAACGAG	GCACAGCTCG	CGTCGCGCAT	GCTGCTGCAG	GTCCACGACG	900
AGCTGCTGTT	CGAAATCGCC	CCCGGTGAAC	GCGAGCGGCT	CGAGGCGCTG	GTGCGCGACA	960
AGATGGGGCG	CGCTTACCGC	CTCGACCTCC	CGCTGAGGCT	GTGCGTGGGC	TACCGCGGCA	1020
GCTGGGACGC	GCGGCGGCAC	TGAGTGGGCA	GCTGCTATCT	GCGGCGGGGA	TTGCGCGATT	1080
TTTCCGCGCT	GAGTTACGCG	TCGCGCGCAT	CGGGACCGAG	TTTGTCCAGC	GTGTACCGCT	1140
CGAGTAGGCT	CCTCA					1155

(2) INFORMATION FOR SEQ ID NO:13:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1771 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCGCCGTC	TGGTGTTTGA	ACGTTTTTAC	CGGTCCGGCAT	CGGCACGGGC	GTTGCCGGGT	60
TCGGGCGCTG	GTTTGGCGAT	CGTCAAAACAG	GTGCTGCTCA	ACCACGGCGG	ATTGCTGCGC	120
ATCGAAGACA	CGGACCCAGG	CGGCGAGGTC	CCTCGAAGCT	CGATTTACCT	GCTGCTCCCC	180
GGCGGTGCGA	TGCGGATTCC	GCAGCTTCCC	GCTGCGACCG	CTGGCGCTCG	GAGCAGCGAC	240
ATCGAAGACT	CTCGGGGTTG	GGCGAACGTT	ATCTCAGTGG	AATCTCAGTC	CACGCGCGCA	300
ACCTAGTTGT	GCAGTTACTG	TTGAAAGCCA	CACCCATGCC	AGTCCACGCA	TGGCCAGATT	360
GGCGCGAGTA	GTGGGCTTAG	TACAGGAAGA	GCAACCTAGC	GACATGACCA	ATCAGCCACG	420
GTATTGCGCA	CGGCGGACG	AGCGGGGAAC	CCGAGGTTAT	GCTCAGGGGC	AGCAGCAAAC	480
GTACAGCGAG	CAGTTGAGCT	GGCCTTACCC	ACGTTCCCGG	CCCGCGCAGC	CAAGCCAGTA	540
CGTCAACCC	TACGAGGCGT	TGGTGGTAC	CGGCGCGGCT	CTGATACCTG	GCTGATTTCC	600

GACCATGACG	CCCCCTCCTG	GGATGCTTGG	CCACGCCCCC	CGTGCAAGCA	TGTTGGCCAT	560
CGGCGCGGTG	ACGATAGGCG	TGGTGTCCGC	CGGCATCGGC	GGCGCGGCGG	CATCCCTGGT	720
CGGCTTCAAC	CGGGCACCCG	CGGGCCCCAG	CGGCGGCCCA	GTGGCTGCCA	GGCGCGCGCC	780
AAGCATCCCC	GCAGCAACAA	TCCCGCCCGG	GTGGGTGAA	CAGGTGGCGG	CCAAGGTGGT	840
GCCCCAGTCT	GTCTGTCTGG	AAACCGATCT	GGGCGGCCAG	TCCGAGGAGG	GCTCCGGCAT	900
CATTCTGTCT	GCCGAGGGGC	TGATCTTGAC	CAACAACAC	GTGATCGCGG	CGGCCGCCAA	960
GCCTCCCTTG	GCCAGTCCGC	CGCGGAAAC	GACGGTAACC	TTCTCTGACG	GGCGGACCGC	1020
ACCCCTCAGG	GTGGTGGGGG	CTGACCCAC	CAGTGATATC	GGCGTGGTCC	GTGTTGAGGG	1080
CGTCTCCGGG	CTCACCCCGA	TCTCCCTGGG	TTCTCTCTCG	GACCTGAGGG	TGGGTGAGCC	1140
GGTGTCTGGG	ATCGGGTGGC	CGCTGGGTTT	GGAGGGGACC	GTGACCAAGG	GGATGCTCAG	1200
CGCTCTCAAC	CGTCCAGTGT	CGACGACCGG	CGAGGCCGGC	AACCAAGACA	CGGTGCTGGA	1260
CGCCATTTCAG	ACCGACGGCG	CGATCAACCC	CGGTAACTCC	GGGGGCGCGC	TGGTGAACAT	1320
GAACGCTCAA	CTCGTGGAG	TCAACTCGGC	CATTGCCACG	CTGGGCGCGG	ACTCAGCCCA	1380
TGCGCAGAGC	GGCTCGATCG	GTCTGGGTTT	TGCTATTCCA	GTGAGCAGG	CCAAGCGCAT	1440
CGCGGACGAG	TTGATCAGCA	CGGGCAAGGC	GTACATGCC	TCCCTGGGTG	TGCAGGTGAC	1500
CAATGACAAA	GACACCCCGG	GCGCCAGAT	CGTCAAGTA	GTGGCCGGTG	GTGCTGCCGC	1560
GAACGCTGGA	GTGCGAAGG	CGGTGGTTGT	CACCAAGGTC	GACGACCGCC	CGATCAACAG	1620
CGCGGACGCG	TTGGTTGGCG	CGGTGCGGTC	CAAGCGCCCG	GGCGCCACGG	TGGCGCTAAC	1680
CTTTCAGGAT	CGCTCGGGCG	GTAGCCGCAC	AGTCCAACTC	ACCTTCGGCA	AGGCGGAGCA	1740
GTGATGAGGG	TCCGCGCCCA	GTGTTCAAGG	C			1771

(2) INFORMATION FOR SEQ ID NO:14:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1858 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCCACCGCG	GTGGCGGCGG	CTCTAGAAGT	AGTGGATCCC	CGGGGCTGCA	GGAAATCGGC	60
ACGAGGATCC	GACGTGCGAG	GTGTGCGAAG	CGCGCGCCGC	GGAGTATCG	GTCCATGCCCT	120
AGCCCGCGGA	CGGCGACCGC	CGGAATGGCG	CGAGTGAGGA	GGCGGGCAAT	TGGCGGGGCG	180
CGGGCGACCG	CGAGCGCCCG	AATGGCGCGA	GTGAGGAGGC	GGCAGTCAAT	GCCCAGCGTG	240
ATCGAATCAA	CCTGCATTGG	GCTGCGGGCG	CCATTGACA	ATCGAGGTAG	TGAGCGCAAA	300
TGAATGATGG	AAAACCGGCG	GTGAGGTCGG	CTGTTCTGGT	GGTGCTAGGT	GCTGCGCTGG	360
CGTTGTGGCT	ATCAGGATGT	TCTTCCCGCA	AACCTGATGC	CGAGGAACAG	GGTGTTCCTG	420
TGAGCCCGAC	GGCTTCGGAC	CCCGGCTTCC	TGCGCGAGAT	CAGGCACTCG	CTTGATCCGA	480
CAAAAGGGTT	GACCAAGCGT	CACGTAGCGG	TCCGAACAC	CGGGAAGTCC	GACAGCTTGC	540
TGGTATTAC	CAGTCCCGAT	GTGACGCTCC	GGGCCAATCC	GCTCGCGGCA	AAGGGCCTAT	600
GCACCTACAA	CGACGAGCAG	GCTGTCCCGT	TTGGGGTACA	AGGCGACAAC	ATCTCGGTGA	660
AACGTTCGGA	CGACTGGAGC	AATCTCGGCT	CGATTTCTGA	ACTGTCAACT	TCAAGCGTGC	720
TGGATCTTGC	CGCTGGGGTG	ACGCAAGTGC	TGTCCGCTGT	CACGAACCTC	CAAGCGCAAG	780
GTACCGAAGT	GATGAGCGGA	ATTTGACCA	CCAAAATCAC	CGGGACCATC	CCCGCGAGCT	840
CTGTCAAGAT	GCTTGATCCT	GGCGCCAGGA	GTGCAAGGCC	GGCGACCGTG	TGGATTGCCC	900
AGGACGGCTC	GCACCACTTC	GTCCGAGCGA	GCATCGACCT	CGGATCCGGG	TGGATTGAGC	960
TCACGCGAGT	GAAATGGGAA	GAACCGGTCA	ACGTGCACTA	GGCGGAAGTT	GCGTGGAGGC	1020
GTTGNTCGAA	ACGCGCTTGT	GAACGGTCTC	AACGGNAC			1058

(2) INFORMATION FOR SEQ ID NO:15:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAATTCGGCA	CGAGAGGTGA	TGGACATCAT	CGGGACCAGC	CCCACATCCT	GGGAACAGGC	60
GGCGGCGGAG	GGGTCCAGC	GGGCGCGGGA	TAGCGTCGAT	GACATCCGCG	TGCTCGGCT	120
CATTGAGCAG	GACATGGCG	TGGACAGCGC	CGGCAAGATC	ACCTACCGCA	TCAAGCTCGA	180
AGTGTCTTTC	AAGATGAGGC	CGGCGCAACC	GCCTAGCAC	GGCGCGGCGA	GCAAGACGCA	240
AAATCGCAGG	GTTCGCGTT	GATTCTGCG	ATTTTGTGTC	TGCTCGCGGA	GGCTTACCAG	300
GGCGCGGCCA	GGTCGCGTG	CTGCGGTATC	CAGGCGTGCA	TGCGGATTCC	GGCGGCCAGC	360
CCGGAGTTAA	TGCTTCGCT	CGACCGGAAC	TGGCGGATCC	GGCGGNGAGC	TGATCGATGA	420
CGGTGGCCAG	CCGCTCGATG	CCCGAGTTGC	CCGAGGAAAC	GTGCTGCCAG	GGCGGTAGGA	480
AGGTCCGCTA	GGCGCGGCTG	CTGACCGGCT	CTGCTTGGCG	CCTCAGTGGC	GGCAGCGAGC	540
GG						542

(2) INFORMATION FOR SEQ ID NO:16:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 913 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGGTGCGGCG	CGCGGCTCCG	TTGCGCCCAT	TGCGCGGCTC	CGCGATCAGC	TGCGCATGCG	60
CACCATCAGC	GGTTTTCCCG	CGCGCAGCGC	CGGTGGCGCC	GGGCGCGCGG	ATGCCACCGC	120
TTGACCTCTG	CGCGCGGCGC	CGGCATTGCC	ATACAGCACC	CGCGCGGGGG	CACCGTTACC	180
GGCGTGCGCA	CGGTCCCGCG	CGGTGCGGTT	TCAGGCGGGG	GAGGCGGAAAT	GAACCGCGCGC	240
CAAGCGCGCG	GGCGGACCGG	TTGCGCGGTT	TTGCGCGGCG	CGCGCGCGCG	CGCGCAATTG	300
CGGAACAGCT	AGGACCGGTT	CGCGCGAGCG	CGCGCGCGCT	TAACCGCGCT	CGCGCGCGCG	360
GGCGCGCGAC	CGCGATTAC	CGCGTTGCG	GTTCGCTGCG	CGCGCGTTAC	CGCGCGCGCG	420
GTTCGCGCGC	AATATTGCG	GGCGACCGCG	AGACCGCGCG	GGCGCGCGAT	TGCGCGCGCG	480
CACCGAAGCA	ACAGGCGAAC	GGTGGCGCGC	GGCGCGCGCT	TTGCGCGCAT	CACCGCGCAT	540
TCACCGCGAG	CACCGCGGTT	AATGTTTATG	AACCGCGTAC	CGCGAGCGCG	GGCGTTATTC	600
CGCGCGCGCG	GAGGCGGTC	CGCGCGCGCG	CGCGAACCGC	CAAAAGCGCG	GGTTTGGCAC	660
CGCGCGCGCG	GGACCGAGCG	GTGCGCGCGA	TGCGCGGTTT	GGCGCGGTTG	CGCGCGCGAT	720
TGCTGCTGCT	GAGCGGTTA	GGCGCGGTTT	CGCGGTTTCC	GGCGGTGGCG	CGTGGCGCGC	780
CGCGCGCGCG	GTTCGCTTAC	AGCGACCGCG	CGGTGGCGCG	GTTCGCGCGA	TTGCGCGCAT	840
TGCGCGGCTT	GGCGCATTC	CGCGGTTTCC	CGCGCGCGAC	CGCGGTTTGG	CGCGCGCGCG	900
CGCGCGCGCG	CGC					913

(2) INFORMATION FOR SEQ ID NO:17:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1872 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GACTACGTTG	GTGTAGAAAA	ATGCTGCGCG	CGGACCGCTT	AAGGCTGGGA	CAATTTCTGA	60
TAGCTACCGC	GACACAGGAG	GTACCGCGAT	GAGCAATTGC	CGCGCGCGCT	CATCAGGCTG	120

GTGATGGTTC	CTGAGCGTTC	TGGCTGCGGT	CGGGCTGGGC	CTGGGCACGG	CGCGGGCCCA	180
GGCGGCCCCG	CGGCGCTTGT	CGCAGGACCG	GTTCGCGGAC	TTCCCCGGCG	TGCCCCGCGA	240
CCCGTCGCGG	ATGGTCGCGC	AAGTGGCGCC	ACAGGTGGTC	AACATCAACA	CCAAACTGGG	300
CTACAACAAC	GGCGTGGGCG	CCGGGACCGG	CATCGTCATC	GATCCCAACG	GTGTGCTGCT	360
GACCAACAAC	CACCTGATCG	CGGGCGCCAC	CGACATCAAT	GGTTTCAGCG	TGCGCTCGCG	420
CCAAACCTAC	GGCTTCGATG	TGGTGGGGTA	TGACCGGACC	CAGGATGTGG	CGGTGCTGCA	480
GCTGCGCGGT	GGCGGTGGCC	TGCGGTGGCG	GGCGATCGGT	GGCGGGCGTG	CGGTTCGTGA	540
GGCGGTGCTC	GCGATGGGCA	ACAGCGGTGG	GCAGGGCGGA	ACGCCCCGTG	CGGTGCTTGG	600
CAGGCTGGTC	GGCTTCGCGC	AAACCGTGCA	GGCGTGGGAT	TGCTTGACCG	GTGCGCAAGA	660
GACATTGAAC	GGTTTGATCC	AGTTGATGCG	CGCAATCCAG	CCCGGTGATT	CGGGCGGGCC	720
CGTCTCAAC	GGCTTAGGAC	AGGTGGTGGG	TATGAACAAG	GGCGCGTCCG	ATAACTTCCA	780
GCTGTCCGAG	GTTGGGCAGG	GATTGCGCAT	TCCGATCGGG	CAGGCGATGG	CGATGCGCGG	840
CCAAATCCGA	TGCGGTGGGG	GGTCACCCAC	CGTTCAATATC	GGGCTTACCG	CGTTGCTCGG	900
CTTGGGTGTT	GTGACAACA	ACGGCAACCG	CGACCGATTC	CAACGGGTGG	TGGAAGCGCG	960
TGCGGCGGCA	AGTCTCGGCA	TCTCCACCGG	CGACGTGATC	ACCGCGGTCC	ACGGCGCTTC	1020
GATCAACTCG	GGCAGCGGCA	TGGCGGACCG	GCTTAAACGG	CATCATCCCG	GTGAAGTCAT	1080
CTCGGTGAAC	TGCAAAACCA	AGTGGGGCGG	CACCGGTACA	GGGAACGTGA	CATTGGGCGA	1140
GGGACCGCGG	GCTGTGATTG	TGCGGATATC	CACCGGCGCG	CGGGCCAATT	GGATTGGCGG	1200
CAGCGGTGAT	TGCGCGGTGA	GGCGCGGAGT	TGCGGTCTCC	GTGCGGTGGG	CATTGTGGAA	1260
GCAATGAACG	AGGAGAGACA	CAGCGTTGAG	CACCGTCCCG	TGCGGGCGAG	TTAGGTGGAA	1320
GGCGGTGTGG	TGAGGATCC	GATGCGCAAG	GACTTGCGCA	GGCGCGCGCG	CCTGCGCGCG	1380
GATCGGACCT	GTTTAAAGCA	CGCGGTCTTC	TACGAGGTGC	TGGTCCGGCG	GTTCTTCGAC	1440
GGCAGCGGCG	ACGCTTCCGN	CGATCTCGGT	GGACTCATCG	ATCGGCTCGA	CTACCTGCGG	1500
TGGGTGGGCA	TGAGTGCAT	CTGTTGCGCG	CGTTCTTACG	ACTCACCGCT	CGCGGACCGC	1560
GTTTACGACA	TTGCGGACTT	CTACAAGGTC	CTGCGCGAAT	TGCGGACGCT	CGACGATTTT	1620
GTGCGCGTGG	TGACACCGCG	TCACCGGCGA	GATATCGGCA	TCATCACCGA	CCTGCTGATG	1680
AATCACACCT	CGGAGTGGCA	CGCGTGGTTT	CAGGAGTCCC	GGCGCGACCG	AGACGAGCGG	1740
TACGGTGAAT	ATTACGTGTC	GAGCGACAGC	AGCGAGCGCT	ACACCGACCG	CGGATCATTC	1800
TTGCTCGACA	CGAAGAGTTC	GAAGTGGTCA	TTGATGCTTG	TCGCGCGACA	GTTCCTACTG	1860
GCACCGATTC	TT					1872

## (2) INFORMATION FOR SEQ ID NO:18:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1482 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTCGCGGAA	ACCTGATGCC	GAGGAACAGG	GTGTTCCCGT	GAGCCCGACG	GGTTCGACCC	60
CGCGCTGCTT	CGCGGAGATC	AGGCACTCGC	TTGATGGGAC	AAAAGGGTTG	ACCAGCGTGC	120
ACGTAGCGGT	CGGAACAACG	GGGAAAGTCC	ACAGCTTGGT	GGTATTACG	AGTCCCGATG	180
TGAGCTGCG	GGGCAATCCG	CTGCGGCGAA	AGGGCGTATG	CACCTACAAC	GACGAGCAGG	240
GTGTCCCGTT	TGCGGTACAA	GGCGACAACA	TCTCGGTCAA	ACTGTTGCGC	GACTGGAGCA	300
ATCTCGGCTC	GATTTCTGAA	CTGTCAACTT	CACCGTGTCT	CGATCTTGCC	GCTGGGGTGA	360
CGCAGCTGCT	GTGCGGTGTC	ACGAACCTTC	AAGCGCAAGG	TACCGAAGTG	ATAGACCGAA	420
TTTCGACCCAC	CAAAATCACC	GGGACCATCC	CGCGGAGCTC	TGTCAAGATG	CTTGATCCTG	480
GGGCAAGAGG	TGCAAGGCGG	GGGACCGTGT	GGATTGCCCC	GGACCGGCTG	CACGACCTCG	540
TGCGAGCGAG	CATCGACCTC	GGATCGGGGT	CGATTGCGCT	CACGCACTCG	AAATGGAAGC	600
AACCGGTCAA	CGTCACTAG	GGCGAAGTTG	CGTGGACCGG	TTGCTCGAAA	CGCCCTTGTC	660
AACGGTGTCA	ACGGCACCGG	AAACTGACCT	CGCTGACCGC	ATCTGAAAAAT	TGACCCCTTA	720
GACCGGGCGG	TGCGTGGTTA	TTCTTCGGTG	GTTCGGGCTG	GTGGGACCGG	GGCGAGGTGC	780
CGGTCTTTGA	GGCGGTAGCT	GTCGCTTTTG	AGGGCGACGA	CTTCAGCATG	GTGACCGAGG	840

CGGTGATCA	TGGCGGCAGC	AACGACGTG	TGCGCGCCGA	AAACCTCGCC	CCACGGGCGG	900
AAGGCTTAT	TGGACGTGAC	GATCAGCTG	GCGGCTCAT	ACCGGGAGGA	CACCAGCTGG	960
AAGAAGAGST	TGGCGGCTC	GGCTCAAAC	GGATGTAAAC	CGACTTCGTC	AACCACCAGG	1020
AGCGGATAGC	GGCCAAACG	GGTGAGTTCG	GGTAGATGC	GGCGGGCTG	GTGAGCTCG	1080
GCGAACCGTG	CTACCCATTC	GGCGGGCTG	GCGAACAGCA	CCCGATGACC	GGCTTGACAC	1140
GCGGTATCG	CCAGGCGGAC	CGCAAGATGA	GTCTTCCCGG	TGCCAGGCGG	GGCCCAAAAA	1200
CACGACGTTA	TGCGGGGCGG	TGATGAAATC	CAGGGTGCCC	AGATGTGCGA	TGGTGTGCGG	1260
TTTGAGGCCA	CGAGCATGCT	CAAAGTCGAA	CTCTTCCAAAC	GACTTCCGAA	CCCGGAAGCG	1320
GGCGGCGCGG	ATGGGGGCGCT	CACCACGTC	GGACTCCCGG	GCTGACACTT	CCCGCTGCAG	1380
GCAGGCGGCG	AGGTATTCTT	CGTGGCTCCA	GTCTCGGCGG	CGGGCGCGAT	CGGCCAGCGG	1440
GGACACTGAC	TCAGGCGAGG	TGGGAGCTTT	CAATGCTCTT	GT		1482

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 876 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAATTCGGCA	CGAGCCGGCG	ATAGCTTCTG	GGCGCGGCGC	GACCAGATGG	CTCGAGGGTT	60
CGTGTCTGGG	GGCAGCGCGG	GGCGCACCAC	CCTGACCGGT	GAGGGGCTTC	AACACGCGCA	120
CGGTCACTCG	TTGCTGCTGG	ACGGCACCAA	CGCGGCGGTG	GTTCCTTACG	ACCGGGCTTT	180
CGCTTACGAA	ATCGGCTACA	TGCGGAAAG	CGGACTGGCC	AGGATGTGGG	GGGAGAACCC	240
GGAGAACATC	TTCTTCTACA	TCACCTCTCA	CAACGAGCGG	TACCTGCAGC	CGCGGAGCGC	300
GGAGAACTTC	GATCCCGAGG	GGTGTCTGGG	GGGTATCTAC	CGTATTCAGG	CGGCCACCGA	360
GCAACGCACC	AACAAGGCGC	AGATCTCTGC	CTCCGGGGTA	GGATGCGCGG	CGCGCTGCGG	420
GGCAGCACAG	ATGCTGCGCG	CGGAGTGGGA	TGTGCGGCGC	GACCTGTGGT	CGGTGACGAG	480
TTGGGCGGAG	CTAAACGCGG	ACCGGGTGGT	CATCGAGACC	GAGAACTTCC	GGCAGCGCGA	540
TGGCGCGCGG	GGGCTGCGCT	ACGTGACGAG	AGCGCTGGAG	AATGCTCGGG	GGCGGCTGAT	600
CGCGGTGTTC	GACTGATGTC	GGCGGCTGCG	CGAGCGGATC	CGACCTGCGG	TGCGGCGCAC	660
ATACCTKAGG	TTGGGCAAGG	ACCGGTTGCG	TTTTTCCGAC	ACTCGGCGCG	CGGCTGCTCG	720
TTACTTCAAC	ACCGACGCGG	AATCCGAGGT	TGCTCGCGGT	TTTGGGAGGG	GTTCGCTGGG	780
TCCAGCGGTC	AATATCGACT	CATTCGCTGC	CGGTCTGCGG	CGGCGCGCGC	AGTTACCGCG	840
ATTTCAGCAA	GGTGGGGGGT	TGCGCGCGAN	TRAGTT			876

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1021 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATCCCGGCGG	GCTGCAGGAA	TTGGGCACGA	GAGACAAAT	TCCAGCGGTT	AATGCAGGAA	60
CAGATTGATA	ACGAATTGAC	ACCGGCACAA	CAATATGTCG	CGATCGCGGT	TTATTTCGAC	120
AGCGAAGACC	TGCGGCASTT	GCGGAGCAT	TTTTACAGCC	AAGCGGTGGA	GGAACGAAAC	180
CATGCAATGA	TGCTGTGCGA	ACGCTGCTC	GACCGGAGCC	TTGTTGTGGA	AATTCCCGGC	240
GAGACACCGG	TGCGAAGCCA	GTTCGACAGA	CCCGCGGAGG	CACTGGGCGT	GGCGCTCGAT	300
CAGGAACCGA	CAGTCACCGA	CCAGGTGCGT	CGGCTGACAG	CGGTGCGCGG	CGACGAGGCG	360
GATTTGCTCG	GCGAGCAGTT	CATGCACTGG	TTCTTGCGAG	AACAGATCGA	AGAGGTGCGC	420

TIGATGGCAA	CCCTGGTGCG	GATTGGCGAT	CGGGCGGGGG	CGAAGCTGTT	CGAGCTAGAG	480
AACCTCGTGG	CACGTGAAGT	GCATGTGGCG	CCGGCCGCAT	CAGGCGCCCC	GCAGCGTGCC	540
GGGGGCCCCC	TCTAGATCCC	TGGGGGGGAT	CAGCGAGTGG	TCCCGTTCCG	CCGCGCGTCT	600
TCCAGGCCAG	CTTGGGTGCG	GCCGGGGTGG	TGAGTACCAA	TCCAGGCCAC	CCCGACCTCC	660
CGGNAAAAGT	CGATGTCTTC	GTACTCATCG	ACGTTCCAGG	AGTACACCGC	CCGGCCCTGA	720
GCTGCCGAGC	GCTCAACGAG	TTGCGGATAT	TCCTTTAAGG	CAGGCAGTGA	GGGTCCACCG	780
GCGGTTGGCC	CGACCGCCGT	GGCGGCACTG	CTGCTCAGGT	ATCGGGGGGT	CTTGGCCAGC	840
AACAACGTCG	GCAGGAGGGG	TGGAGCCCCC	CGGATCCGCA	GACCGGGGGG	GCGAAACGAA	900
CATCAACACC	GCACGGGATC	GATCTGCCGA	GGGGGGTGCG	GGAATACCGA	ACCGGTGTAG	960
GAGCGCCAGC	AGTTGTTTTT	CCACCACCGA	AGCGTTTTCC	GGTCATCGGN	GGCMTTAAAG	1020
T						1021

## (3) INFORMATION FOR SEQ ID NO:21:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGTGCCGAGC	AACGGAAGAA	CACAACCATG	AAGATGGTGA	AATCGATCGC	CGCAGGTCTG	60
ACCGCCGCGG	CTGCAATCGG	CGCCGCTGCG	GCCGGTGTGA	CTTCGATCAT	GGCTGGGGGN	120
CGCGTGGTAT	ACGAGATSCA	GCGGCTCTTC	TTGCGGCGCG	CAGTCCCGTT	GGACCCGONA	180
TCCGCCCCCTG	ANGTCCCSAC	CGCCGCCAG	TGGACCAAGC	TGCTCAACAG	NGTCGNCGAT	240
CGCAACGTGT	CGTTTNGAA	CAAGGAGAT	CTGGTCCAGG	GNGGATCCGG	NGGNAACGAG	300
GCGNGGATC	GNGGATCACA	A				321

## (2) INFORMATION FOR SEQ ID NO:22:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCTTATCGGT	TCCGGTTGGC	GACGGGTTTT	GGNGCGGGGT	GCTTAACCGG	CTCGGCCAGC	60
CGATCGACCG	GCGCGGAGAC	GTGCACTCGG	ATACTCGGCG	CGCGCTGGAG	CTCCAGGCGC	120
CTTCGGTGGT	GNACCGGCAA	GCGGTGAAGG	AGCGTTGNA	GACCGGGATC	AAGGCGATTC	180
ACCGGAGTAC	CCGATCCGGT	CGCGGGCAGC	GCCAGCTGAT	CATCGGGGAC	CGCAAGACCG	240
GCAAAAACCG	CGTCTGTGT	CGGACACCAT	CCTCAACCA	GCGGGAAGAA	CTGGGAGTCC	300
GCTGGATCCC	AAGAAGCAGG	TGCGGTTGTG	TATACGTTGG	CCATCGGGCA	AGAGGGGAA	360
CTTACCATCG	CGG					373

## (2) INFORMATION FOR SEQ ID NO:23:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 352 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GTGACGCGGT	GATGGGATTC	CTGGGCGGGG	CGGCTCGGCT	GGCGGTGGTG	GATCAGCAAC	60
TGGTTACCGG	GGTCCCGCAA	GGCTGGTGGT	TTGCTCAGGC	AGCCGCTGTG	CGGCTGGTGT	120
TCTTGACGGC	CTGGTACGGG	TTGGCCGATT	TAGCCGAAAT	CAAGGCGGGC	GAATCGGTGC	180
TGATCCATGC	CGGTACCGGC	GGTGTGGGCA	TGGCGGCTGT	GCAGCTGGCT	CGCCAGTGGG	240
GGGTGGAGGT	TTTCGTCAAC	GCCAGCGGTG	GNAAGTGGGA	CACGCTGCGC	GCCATNNGT	300
TTGACGACGA	NCCATATCGG	NGATTCCGNC	ACATNCGAAG	TTCCGANGGA	GA	352

## (2) INFORMATION FOR SEQ ID NO:24:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 726 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAAATCGGCG	TTCAATCCGT	TGACGACCG	GCTGGCGATA	ATCGACGAAG	TGATCAAGCC	60
GGGTTTCCCG	GGGCTCATCG	GTACAGCGCA	GTAATCAGCA	ACTTCTCTGG	TATATCGCAC	120
CTAGCGTCCA	GTTCCTTGGC	AGATGGCTTT	CGTACCGTCA	TGGCATGTAC	CGGTTCGGCT	180
GGCGCACGCT	CATGCTGGCG	GGTTCATCC	TGGGACGGGG	TGTGGCGGGT	CTCGGGGTTC	240
GGCGCGAGTC	CGCAGCCCAA	ACCGCGCCCG	TGCGCGACTA	CTACTGGTGC	CGCGGGCAGC	300
CTTTGAGCCG	CGCATGGGGG	CGCACTGGGG	ATCCCTACAC	CTGCCATGAC	GACTTTCACC	360
GGGACAGCGA	CGGCCCCGAC	CACAGCGCGG	ACTACCCCGG	ACCCATCCTC	GAAGGTCCCG	420
TGCTTGACGA	TCCCGGTGGT	GGGCGCGGGC	CGCGGCTGTC	CGGTGGCGGC	GCATAGCGCT	480
CGTTGACCGG	GGGCGATCGG	CGAATACCGG	TATAAAGCCG	GGGCTGCGCC	CGGCAAGCTA	540
CGACCCCGCG	CGGGGCGAGT	TTACGCTCCC	GTGCGGATGG	ATCGCGCGGT	CGGATGACAG	600
AAAATAGCGG	ACGGTTTTTG	CAACCGCTTG	GAGGACGCTT	GAAGGGAACC	TGTCTGTAAC	660
GGCGACAGCG	CGTCCACCAT	CGACATCGAC	AAGGTTGTTA	CCGCGACACC	CGTTCCCGGG	720
ATCGTG						726

## (2) INFORMATION FOR SEQ ID NO:25:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 580 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCGACGAGG	ACGAACGTGG	GGCCACCCAC	CGGCTATGCG	TTGATGCAAG	CGACCCGGAT	60
GGTCCGCGAC	CATATCCAAAG	CATGCTGGGT	GGCCACTGAG	CGACCTTTTG	ACCAGCCCGG	120
CTGCCCGATG	GGGCCCCGGT	GAAGTCATTG	CGCCGGGGCT	TGTGCACCTG	ATGAACCCGA	180
ATAGGGAACA	ATAGGGGGGT	GATTTGGCAG	TTCAATGTGG	GGTATGGCTG	GAAATCCAAT	240
GGCGGGGCAT	GCTCGGGGCC	GACCAAGGCTC	GGCGAGGCGG	GCCAGCCCGA	ATCTGGAGGG	300
AGCACTCAAT	GGCGGGCATG	AAGCCCCGGA	CGGGCGAGCG	TCCTTTGGAA	GCACTAAGG	360
AAGGGCGCGG	CATTGTGATG	CGAGTACCAC	TTGAGGGTGG	CGGTGCGCTG	GTCGTGAGGC	420
TGACACCCGA	CGAAGCCGCG	GCACTGGGTG	ACGAATCAA	AGGCGTTACT	AGCTAAGACC	480
AGCCCAACGG	CGAATGGTGG	GCCTTACCGG	CACACCTTCC	GGTAGATGTC	CAGTGTCTGC	540
TCGGCGATGT	ATGCCCCAGGA	GAACTCTTGG	ATACAGCGCT			580

## (2) INFORMATION FOR SEQ ID NO:26:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

AACGGAGGCG CGGGGGTTT TGGCGGGGCT GGGGCGGTG GCGGCAACGG CGGGGCGGGC    60
GGTACCGCGG GGTGTTCGG TGTGGCGGG GCGGTTGGG CCGGAGGCAA CGGCATCGCC    120
GGTGTACAGG GTACGTGGG CAGCACACCG GGTGGATCCG                    160

```

## (2) INFORMATION FOR SEQ ID NO:27:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

GACACCGATA CGATGCTGAT GTACGCCAAC GTTGTGACAA CGTTCGAGGC GTTCACGATC    60
CAGCGCACAC CCGACGGGCT GACCATGGGC GATCGGCCCC CTTTCGCGGA GCGGGCTGCC    120
AAGGCGATGG GAATCGACAA GCTGCGGGTA ATTCTACCG GAATCGACCC CTTCTCGGCT    180
GAACGCGAAC AGTGGGACGA CCGCAACAAC ACCTTGGCGT TGGCGCCCCG TGTGTTTCTC    240
GCCTACGAGC GCAACGTACA GACCAACGCC CG                    272

```

## (2) INFORMATION FOR SEQ ID NO:28:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 317 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```

GCAGCCGGTG GTTCTCGGAC TATCTGGGCA CGGTGACGCA GCGCGACGTG CCGGAGCTGA    60
AGCGGATCGA GCGACGGGAT CGCTGCGGCG GGTTCATGCC CTACCTGGCC GGTATCAGCG    120
CGCAGGAGCT GAACCTGGCC GAACCGGCGC GGGTCATCGG GGTGACGCG GCGACGATCC    180
GTTGGGATCT GCGGTGGTTC GAGACGGTCT ATCTGGTACA TCGGTTGCCC GCTTGGTGGC    240
GGAATCTGAC CCGAAGATC AGAAGCGGT CRAAGATCCA CTTGTGCGAC AGTGGCTTCG    300
CGGCTGTGTT GCGCGGCG                    317

```

## (2) INFORMATION FOR SEQ ID NO:29:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCGTGGAG	CTGTCTATGA	ACAGCTTTGC	CGGACGCGCG	GCGGCCAGCA	CGTCTGTGTA	50
GCAGCGCGCG	ACCAGCTTCC	CGGTGGGCGG	CATGGTGATG	ACCAGCTCGG	CCTCGGCCAC	120
CGCTTCGGCG	GCGCTACGAA	ACACCGCGAC	ACCCTGCGCG	GCGCGCGCGG	ACGCGCGCGT	180
GG						182

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GATCGCGAAG	TTTGGTGAAG	AGGTGGTCTG	CGCGAAGCTC	TGGCGCGCTG	CGAAGCGGGT	50
CGGCGTTTCA	GAGGCGAAGA	CACGCTTCTC	CGAGCTGCTG	CGGCTCTCTT	ACGCGCGGCA	120
GAGGTTGAGA	TTGCGCGCGG	CGGCGAGCGG	GTAGCAAGGC	TTGTGCGGCT	GCATCTTCAT	180
GAGACTCGGC	GTTTAGGCAT	TCACCATGGC	GTGTACGCGG	TGCGCGAGCA	TTTGGAGGCT	240
CGTTTGTGAG	ACGAGCTGCT	CGAAGCTTTT	CACCGGTGAA	GCGCTACCTC	ATCGACACCC	300
ACGTTTGG						308

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 267 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGGACGACCA	GCAACTCAGC	TGGATGATCG	TGGCGAGCGG	CATTGAGGAC	GGAGAGAATC	50
CGGCGAAGGC	TGGCGCGCGG	CAAGTCTCTA	TAGTGACCGG	CGGTAGAGGG	CTCCCCCGAT	120
GGCAGCGGAC	TATTCTGCTG	TGCGCTCTGC	CGGTAAAGGC	GGGTAAAGAA	ATGTGAGCGG	180
ACACGATGAG	CAATCAGACC	TACCGAGTGA	TGGAGATCTT	CGGGACCTCG	CTCGAGCGCG	240
TGGAGCGCGG	AATCCAGGCG	GATCTGG				267

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1515 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTCGTCCCGA	AAGAATGTGA	GGGACACGAA	TGAGCAATCA	CACCTACCGA	GTGATCGAGA	50
TGCTCGGGAC	CTCGCCCGAC	GGCTCGACG	CGGCAATCCA	GCGCGGTCTG	GCCGAGCTG	120
CGCAGACCAT	GCGCGCGCTG	GACTGGTTTG	AAGTACAGTC	AATTGAGGGC	CACCTGGTGG	180
ACGGAGCGGT	CGCGCACTTC	CAGGTGACTA	TGAAAGTGGG	CTTCGCGCTGG	AGGATTCCTG	240
AACCTTCAGG	CGCGCGCGAT	AAGTGAGGTG	CATCATTAAAG	CGACTTTTCC	AGAACATCCT	300
GACGCGCTCG	AAACGCGGTT	CAGCGGACGG	TGGCTCCGCG	GAGGCGCTGG	CTCCAAAATC	360
CCTGCGACAA	TTGCTCGCGG	GCGCTACAAA	GGAAGTGGGT	GCTGAATTGG	TGGGTTATCT	420

GCTCGACCTG	TGTGGGCTGC	AGCCGACCGA	AGCGGTGCTC	GACGTGGGCT	GCGGCTCGGG	480
GCGGATGGCG	TTGCCCTCA	CCGGCTATCT	GAACAGCGAG	GGACGCTACG	CCGGCTTCCA	540
TATCTCGCAG	AAAGCCATCG	CGTGGTGCCA	GGAGCAGATC	ACCTCGGCGC	ACCCCAACTT	600
CCAGTTCCAG	GTCTCCGACA	TCTACAACTC	GCTGTACAAC	CCGAAAGGGA	AATACCACTC	660
ACTAGACTTT	CGCTTTCCAT	ATCCGATGTC	GTCTTTGGAT	GTGGTGTTC	TTACCTCGGT	720
GTTCAACCCAC	ATGTTTCCGC	CGGACGTGGA	GCCTATCTG	GACGAGATCT	CCCGCGTCTT	780
GAAGCCCGGC	GGACGATGCC	TGTGACGTA	CTTCTTGCTC	AATGACGAGT	CGTTAGCCCA	840
CATCGCGGAA	GGAAAGAGTG	CGCACAACTT	CCAGCATGAG	GGACCGGCTT	ATCGGACAAT	900
CCACAGAGAG	CGCCCGGAG	AAGCAATCGG	CTTCCCGGAG	ACCTTCCTCA	GGGATGTCTA	960
TGGCAAGTTC	GGCTTCGGCG	TGLACGAACC	ATTGCACTAC	GGCTCATGGA	GTGGCCCGGA	1020
ACCACGCTTA	AGCTTCCAGG	ACATCTTCAT	CGCGACCAAA	ACCGCGAGCT	AGGTCCGCAT	1080
CCGGGAAGCA	TCCGACACCC	GTGGCGCCGA	GCGCGGCTGC	CGGCAGGCGG	ATTAGGCGGG	1140
CAGATTAGCC	CGCGCGGCT	CCCGGCTCCG	AGTACCGGCC	CCCGAATGGC	GTCAACCGCT	1200
GGTAACCCAG	CTTCCGCGCC	TGGCGCGCGG	CGTCCCGGAT	CAGGTGGTAG	ATCCCGACAA	1260
AGCTTCGCTG	ATCGGTCTATC	ACCAACGGTG	ACACGAGCTG	GTGTGTCACC	AGCCCGAAGC	1320
CCACCCCGGT	CTCCGGTCTT	GTCCAGCCGA	TGAGCCGCGC	CAAGCCCACT	TGACCAAAACC	1380
CCGCGATCAC	GTTCGCGATC	GGCATACCTT	GATAGCCCAAG	ATGAAAATTT	AAGGCGACCA	1440
ATAGATTTCC	ATCCGCGAGA	ACTTCCGCTC	GCTTCCGCGT	CAGGCGCGTG	ACCAGCTCCC	1500
CGGACAAGAA	CGGTATGCGG	TGATCTCGCC	CTCGTCCCG			1539

(2) INFORMATION FOR SEQ ID NO:33:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 851 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTCCAGGGTG	GCTGATGA	GCTCACCGC	GGGCGAGGCG	GAGCTGACCG	CGCCCCAGGT	60
CGGGTTGCT	CGGGCGGCTT	ACGAGACCGC	GTATGGGCTG	ACGGTGGCCC	CGCCGCTGAT	120
CGCCGAGAAC	CGTCTGAAC	TGATGATTCT	GATAGCGACC	AACCTTTTGG	GGCAAAACAC	180
CCCGCGGATC	CGGTCAACG	AGCCGGAATA	CGCGGAGATG	TGGGCCCCAG	ACGCGCGCCG	240
GATGTTTGGC	TACGCGCGCG	CGACCGGAGC	GGCGACGGCG	ACGTTGCTGC	CGTTCCAGGA	300
GGCGCGGGAG	ATGACCAAGC	CGGGTGGGCT	CTTCGAGCAG	CGCCCGCGCG	TGAGGAGGCG	360
CTCCGACACT	CGCGCGCGCA	ACCAGTTGAT	GAACATGTC	CCCCAGGCGC	TGAAACAGTT	420
GGCTGAGCCC	ACCGAGGCGA	CCACGCGCTG	TTCCAGGCTG	GCTGGGCTGT	GGAGACCGGT	480
CTCCCGGAT	CGGTCCCGCA	TCAGCAACAT	GCTGTGATG	GCCAACAACC	ACATGTGAT	540
GACCAACTCG	GCTGTGTCGA	TGACCAACAC	CTTGAGCTCG	ATGTTGAGGG	GCTTTGCTCC	600
GGCGCGCGCG	CGCCAGGCGG	TGCAAAACCG	GGCCAAAAC	GGGCTCGCGG	CGATGAGCTC	660
GCTGGGCGAG	TGCTGGGTT	CTTCGGGCTT	GGCGGCTGGG	GTGGCGCGCA	ACTTGGGTCC	720
GGCGCGCTCG	GTACGCTATG	GTCACCGGGA	TGGCGGAAA	TATGCAAGT	CTGGTCGGCG	780
GAACGCTGGT	CCGGCTTAAG	GTTTACCCCC	GTTTTCTGGA	TGCGGTGAAC	TTGCTCAACG	840
GAAACAGTTA	C					851

(2) INFORMATION FOR SEQ ID NO:34:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 254 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATCGATCGG	GCGGAAATTT	GGACCAAGATT	CGCCTCCGGC	GATAACCCAA	TCAATCCAAAC	50
CTAGATTTAT	TGGTCCAGG	GGCCCGAGTA	ATGGCTCGCA	GGAGAGGAAC	CTTACTGCTG	120
CGGGACCTG	TGGTAGGTC	TGATACGGC	GGAAGGCGTC	GACATTTTCC	ACCGACACTC	180
CCATCCAAAC	GTTCGAGGGC	CACTCCAGCT	TGTAGCGAG	GCGACGCACT	CGCAGGCTTC	240
GCTTGGTCAA	GATC					284

## (3) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1227 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GATCTTGACC	GAAGCTGGGG	CGGCCAAGGC	GAAGTGGCTG	TTGGACCCAG	AGGGACGGGA	50
CGATCTGGCG	CTGGGATGG	CGGTCAGCG	GGGGGGGTC	GCTGGATTGC	GCTATAACCT	120
TTTCTTGGAC	GACCGGACGC	TGATGGTGA	CCAAACCGCG	GAGTTGGCTG	GTGTCAGGTT	180
GATCGTGGAC	CGGATGAGCG	CGCCGTATGT	GGAAGGCGCC	TGGATCGATT	TGCTCGACAC	240
TATGAGAAAG	CAAGGTTCC	CATCGACAAT	CCCAACGCCA	CGGGCTCGTG	CGCGTGCGGG	300
GATTCGTTCA	ACTGATAAAA	CGCTAGTAGG	ACCCCGCGGT	GCGCAACACG	TACGAGCACA	360
CCAGACCTG	ACCGCGCTGG	AAAAGCAACT	GAGCGATGCG	TTCGACCTGA	CGCGGTGGCG	420
GGCGCGCGCG	GGCAGGTGTC	ACCTGCATGG	TGAACAGCAC	CTGGGCTTGA	TATTGGGACT	480
AGTACACGAT	TTTGTGATC	GAGGTCACTT	CGACCTGGGA	GAACTGCTTG	CGGAACCGCT	540
CGCTGCTCAG	CTTGGGCCAAG	GCCTGATCGG	AGCGCTTGTC	GCGCACGCGG	TGCTGGATAC	600
CGCACAGCGC	ATTGCGAAGC	ATGGGTGTCA	CATCGCGGTT	CTCCAGCGCG	TTGAGGTATC	660
CCTGAATGCG	GGTTTTGGCC	GGTCCCTCGG	AGAATGTGCC	TGCGGTGTTG	GCTCCGTTGG	720
TGCGGACCGC	GTATATGATC	GCGCGCGTCA	TAGCGGACAC	CAGCGCGAGG	GCTACCACAA	780
TGCGGATCAG	CAGCGGCTTG	TGCGGTCCCT	TGCGGTAGGA	CACCTGCGGC	GGCAGCGCGG	840
GATATGCGGC	GGCGGGCAGC	GCGCGGTGCT	CTGCGGCTCC	CGGGGCGAAG	GCGGTTGCGG	900
CGCGCGCGAG	GTCTGGGGG	TAGTCCAGGG	CTTGGGGTTC	GTGGGATGAG	GGCTCGGGGT	960
ACCGCGCGCG	TGCGTTGGTG	CGGACACCGG	GCTTCGGCGA	GTGGGACCGG	GGCAATGTGG	1020
TTCTCTTAGG	GTGGTGGAGC	GGACGAGCTG	CTAGGGGCGC	AACCGCGCGT	CGCGTCAAGC	1080
GGCAGCATCG	GCAATCAGGT	GAGTCCCTTA	GCGAGGCTAG	CGCAACAGCT	GCGGTCAAGT	1140
CTCAAGCGGA	CGGGGCGCGG	CGCGCGCGCG	ATAATGTGGA	AAGACTAGGC	AACCTTAGGA	1200
ACGAAGGAGC	GAGATTTTTT	GACGATC				1227

## (2) INFORMATION FOR SEQ ID NO:36:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CGCGTGTGGG	CGGATCGGGC	GGGTGGTTGA	ACGGCAACCG	CGGGGCGCGC	GGGGCGGGCG	50
GGACCGGCGC	TAACGGTGCT	GCCGGCGGCA	ACGCGTGGTT	GTTCGGGGCT	GGGGGCTCGG	120
GCGGCGCGCG	CACCAATGCT	GCGTGGCGCG	GCTCCGCGCG	ATTGTCTTAC	GGCAACGGCG	180
G						181

## (2) INFORMATION FOR SEQ ID NO:37:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GCGGTGTCCG	CGGATCCGGC	GGGTGCTTGA	ACGGCAACGG	CGGTGTCCGC	GGCCGGGGCG	60
GCGACGGCGT	CTTTGCCGGT	GCCGGCGGCC	AGGGCGGGCT	CGGTGGGCAG	GGCGGCATG	120
CGGGCGGCTC	CACCGGCGGC	AACGGCGGTC	TTGGCGGGCG	GGCGGTGGC	GGAGGCACG	180
CGCCGACCG	CGGCTTCGGT	GGCAACGGCG	GTAAGGGTGG	CGAGGGCGGN	ATTGGCGGCG	240
GCATTCAGAG	CGGACCGGC	CTCGGCGGTG	ACGGCGGTGA	CGCGGTGAC		290

## (2) INFORMATION FOR SEQ ID NO:38:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCACTGG	CATGCGGGT	GTCAGTGGAA	GCAT	34
-----------	-----------	------------	------	----

## (2) INFORMATION FOR SEQ ID NO:39:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GATGCTGCT	CGTCCCCCG	TTGGCGCGA	CGCCACCGGT	CCACCGTTA	CGGAACAAGC	60
TGGGTGGTC	GCCAGCACCC	CGGACCGGC	CGACCGCGGA	GTGGAACAT	GGACCGGTCC	120
TATCCCCACC	ATTGCGGCG	GNCCACCGG	CACCG			155

## (2) INFORMATION FOR SEQ ID NO:40:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGGCGTTCA	CGGGCGGCG	GGACCGGGC	AGCCCCGNGG	GGCCCGGGGG	TGG	53
------------	-----------	-----------	------------	------------	-----	----

## (2) INFORMATION FOR SEQ ID NO:41:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GATCCACCGC	GGGTGCAGAT	GGTCCCCGCG	GGCCACCC	GACACGGGC	GGCAACGGCG	60
GCACCGGGCG	CAACGGGCG	AACGCCACCG	TCGTGGGCG	GGCCGGGCG	GGCGGGCGCA	120
AGGGCGGCAA	CG					132

## (2) INFORMATION FOR SEQ ID NO:42:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GATCGGCGGC	CGGNACGGNC	GGGGACGGCG	GGAGGGCGCG	NAACGGGGCG	GGCGNAGCCA	60
CCNGCCCAAGA	ATCCTCGGCG	TCNCCCAATG	GGCGAATGG	CGGACAGGGC	GGCAACGGCG	120
GCANCGGCGG	CA					132

## (2) INFORMATION FOR SEQ ID NO:43:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CGGCACGAGG	ATCGGTACCC	CGCGGCATCG	GCAGCTGCGG	ATTGCGCGGG	TTTCCCCACC	60
CGAGGAAAGC	CGCTACCAGA	TGGCGCTGCC	GAAGTAGGGC	GATCCCTTCG	CGATCCCGGC	120
ATGAACGGGC	GGCATCAAAAT	TAGTGCAGGA	ACCTTTCAGT	TTAGCGACGA	TAATGGCTAT	180
AGCACTAAGG	AGGATGATCC	GATATGACGC	AGTCCGAGAC	CCTGACGGTG	GATCAGCAAG	240
AGATTTTGA	CAGGCGCAAC	GAGGTGGAGC	CCCCGATGGC	GGACCCACCG	ACTGATGTCC	300
CCATCACACC	GTGCGAACTC	ACGGGCGGTA	AAAACGCCGC	CCAACAGNTG	GTNTTGTCCG	360
CCGACAACAT	GCGGGAATAC	CTGGCGGGCG	GTGCCAAAGA	GCGGCAGCGT	CTGGCGACCT	420
CGCTGCGCAA	CGCGGCCAAG	GNGTATGGCG	AGGTTGATGA	GGAGGCTGCG	ACCGGCTTGG	480
ACAACGACCG	CGAAGGAACCT	GTGCAGGCG	AATCGGCGCG	GGCCGTCCGA	GGGGACACTT	540
CGGCGGAACCT	AACCGATACG	CGGAGGCTGG	CCACGGCGCG	TGAACCCAC	ITCATGGATC	600
TCAAAGAAGC	GGCAAGGAAG	CTCGAAACCG	GCGACCAAGG	CGCATCGCTC	GCGCACTGNG	660
GGGATGGGTG	GAACACTTNC	ACCCTGACGC	TGCAAGGCGA	CG		702

## (2) INFORMATION FOR SEQ ID NO:44:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 298 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAAGCCGCGAG	CGCTGTGCGG	CGACGTGGCG	GTCAAAGCGG	CATCGCTCGG	TGGCGGTGGA	60
GGCGCGCGGG	TGCGGTGCGC	GCGGTGCGGA	TGCGCGATCG	GGCGCGCGGA	ATCGGTGCGG	120
CCCGGTGCGG	CTCGGTGACAT	TGCGCGCTTA	GGCGAGGAA	GGCGCGCGCG	CGCGCGCGCG	180
CTCGCGCGCG	GTGGCATGGG	AATGCGGATG	GCTGCGCGCG	ATCGGGGACA	AGCGCGCGCG	240
AAGTCCAAAG	GTCTCTCAGCA	GGAAGACGAG	GCGCTCTACA	CGAGGATCC	TCGTGCGG	300

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CGGCACGAGG	ATCGAATGCG	GTGCGCGGGA	GCACAGCGTC	GCACTGCACC	AGTGGAGGAG	60
CCATGACCTA	CTGCGCGGCT	AACCGCGGAT	ACCGCGCAAGC	GCAGCGCGCA	GGCTCTTACG	120
GAGGCGTCAC	ACCGTCTGTC	GCGGACGCGG	ATGAGGGTGC	GAGCAAGCTA	CCGATGTACC	180
TGAACATGCG	GCTGCGAGTG	CTCGGTCTGG	CTGCGTACTT	CGCGAGCTTC	GGCGCAATGT	240
TCACCGTCAG	TACCGAAGTC	GGCGGGGGTG	ATGCGCGAGT	GTGCGGTGAC	ACTGGGCTGC	300
CGGTGCGGGT	GGCTCTGCTG	GCTGCGGTGC	TTGCGCGGCT	GCTTCTGCTG	CCTAAGGCGA	360
AGAGGATGCT	GAGGTAGATT	GCGGTGCTCG	GCGTACTCGG	CGTATTTCTG	ATGGTCTCGG	420
CGACGTTTAA	CAAGGCCAGC	GCCTATTGGA	CGCGTTGGGC	ATTGTGGGTT	GTGTTGGCTT	480
TCATCTGTTT	CCAGGCGGTT	GCGGCGAGTC	TGCGGCTCTT	GCTGGAGACC	GGCGCTATCA	540
CGCGCGCGGC	CGCGCGCGGC	AAGTTGAGCC	CGTATGAGCA	GTACGCGCGG	TACGGGCACT	600
ACCGGCGAGT	CGGGGTGCGG	ACTACGCTCA	GCAGGGTGCT	CAGCAGCGCG		660
CGGGACTGCA	GTGCGCGCGC	CGCGAGCACT	CTCGCGAGCC	TGCGGATAT	GGGTGCGAGT	720
ACCGCGGCTA	TTGCTGCACT	CGAGGCGAAT	CGGGCGAGTG	ATACACTGCT	CAGCGCGCGG	780
CGGACCGCGC	GGCGAGCTGC	GGGTGCGAAC	AATCGCACCA	GGCGGCAATC	ACGCGACCTA	840
CGGGCTTTCC	GAGGTTGAGC	CGACGAGCAC	CGGTGAGTGC	CGGGACGCGG	TGCGAGGCTG	900
GTTGCGGTCC	AGTCAACTAT	TCAAACCGCA	CGGGGGGGGA	GCAGTCGTCC	TGCGCGGGGG	960
GGCGCGCGGT	CTAACCGCGC	GTTGCGGCGT	CGGTGCGCGC	GTGTGCGCGA	AGAGTGAACA	1020
GGGTGTCAGC	AAGCGCGGAC	GATCTCTGTC	CGGAATTC			1098

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 327 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CGGCACGAGA	GACCGATGCC	GCTACCGTCC	CGCAGGAGGC	AGGTAATTTT	GAGCGGATCT	60
CGGGCGAGCT	GAAAACCGAG	ATCGAGCAGG	TGGAGTCGAC	GGCAGGTTCC	TTGCGAGGCT	120
AGTGGCGCGG	CGCGCGCGGG	ACGCGCGCGC	AGGCGCGCGT	GCTGCGCTTC	CAAGAGGCGG	180
CCAATAAGCA	GAAGCAGGAA	CTCGACGAGA	TCTCGACGAA	TATTCTCTAG	GGCGCGGTCC	240
AATACTCGAG	GGCGGACGAG	GAGCAGCAGC	AGGCGCTGTC	CTCGCLAATG	GGTTCCTGAC	300
CGGCTAATAC	GAAAAGAAAC	GGAGCAA				327

## (2) INFORMATION FOR SEQ ID NO:47:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGGTGCGGAT GATGGCGTTS TCGAACGTGA CCGATTCTGT ACCGCGGTGG	TTGAGATCAA	60
CCAACAACGT GTTGGCGTCS GCAAATGTGC CGAACCCGTG GATCTCGGTG	ATCTTGTCT	120
TTTTCATCAG GAAGTGCACA CCGGCCACCC TGCCCTCGGN	TACCTTTGG	170

## (2) INFORMATION FOR SEQ ID NO:48:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 127 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GATCGGCGG CACGGGGGCT GCGCGCGCA GCACCGCTG CGTGGCGGC	AACGGCGGG	60
CGGGGCTGG CGGCGGAACC GGTGGGTTC TTTCCGCAA CGGCGGTGC	GGCGGCGAC	120
GGGCGCT		127

## (2) INFORMATION FOR SEQ ID NO:49:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CGCGGCGAG GGCGGCGCG CGGCAACGG GAGCGCGCG CGCGCGGCA	ACGCGCGCAA	60
CGCGGCGTCC GGCTTCAAG	G	81

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 149 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GATCAGGCT GCGCGGCTCC GCGCAGAAG GCGTAACGG	AGGAGCTGCC	GGATTGTTT	60
GCAACGGGG GCGCGGCTT GCGCGCGGT CCAACCAAG	CGTAACGGC	GGGCGCGCG	120
GAAACGGTG TGCGGCTGG CTGATTTG			149

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CGGCACGAGA TCACACCTAC CGAGTGATCG AGATCGTCGG GACCTCGCCC GACGTTTTCG	60
ACGCGGHAAT CCAGGGGCGT CTGGCCCGAG CTGCGCAGAC CATGCGCGCG CTGCACTGGT	120
TCGAAGTACA GTCAATTCA GGCACCTCG TCGACCGAGC GGTGCGCGAC TTCCAGGTGA	180
CTATGAAAGT CGGCTTCGCG CTGCAAGGAT CTGGAACCTT CAAGCGCGCG CGATAACTGA	240
GGTGCATCAT TAAGCGACTT TTCCAGAACA TCTGACGCG CTGGAACGCG GGTTCAGCGG	300
ACGCTGCGTC CGCCGAGGCG CTGCTTCGAA AATCGTGGC ACAATTCGTC GCGCG	355

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 999 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATGCATCACC ATCACCATCA CATGCATCAG GTGCAACCCA ACTTGACACG TCGCAAGGGA	60
CGATTGGCGG CACTGGCTAT CGCGGCGATG GCGAGCGCCA GCTGCTGAC GGTTCGGGTG	120
CGCGCGACCG CCAACGCGGA TCGGAGGCG GCGCGCGCGG TACCCACAC GCGCGCTTCG	180
CGCGCGTCCA CGGCTGCAGC GCCACCTCCA GCGCGGACAC CTGTTGCGCG CCGACCAACG	240
GCGCGCGGCA ACACGCGGAA TCGCGAGCGG GCGGATCCCA AGCGAGCACC TCGCGCGCGG	300
GACCGGAAAG CAACGCGCGG ACCTGTCATT GCTGCAAGCG CACCGCAACC TGTCCGGATC	360
GACAACTCGG TTGGAGGATT CAGCTTCGCG CTGCGTGGTG GCTGGGTGGA GTCTGACCGC	420
GCGCACTTCG ACTACGCTTC AGCACTCCTC AGCAAAACCA CCGGGGACCG GCCATTTCGC	480
GGACAGCGCG CGCGGCTGGC CAATGACACC GGTATCGTGC TCGCGCGGCT AGACCAAAAG	540
CTTTACGCGA GCGCGGAAGC CACCGACTCC AAGGCTCGCG CCGGTTGGG CTCGGACATG	600
GGTGAGTTCT ATATGCGCTA CCGGGGACCG CGGATCAACC AGGAAACCTT CTCGCTCGAC	660
GCGAACGCGG TGTCTGGAAG CCGCTCGTAT TACGAAGTCA AGTTCAGCGA TCCGAGTAAG	720
CGGAACGCGC AGATCTGGAC GCGCGTAATC GGTTCGCGCG CCGCGAAGCG ACCGAGCGCG	780
GCGCGCGCTC AGCGCTGGT TGTGCTATGG CTGCGGACCG CCAACAACCG GGTGGACAAG	840
GCGCGCGCGA AGCGCTGGC CGAATCGATC CGGCTTTGG TCGCGCGCGG GCGCGCGCGG	900
GCGCGCGCTC CTGCGAGGCG CGCTCGCGCG CCGCGCGCGG CCGCGGAAGT CGCTCCTACC	960
CGGACGACAC CGACACCGCA GCGGACTTGA CCGGCTTGA	999

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met His His His His His His Met His Gln Val Asp Pro Asn Leu Thr  
 1 5 10 15  
 Arg Arg Lys Gly Arg Leu Ala Ala Leu Ala Ile Ala Ala Met Ala Ser  
 20 25 30  
 Ala Ser Leu Val Thr Val Ala Val Pro Ala Thr Ala Asn Ala Asp Pro  
 35 40 45  
 Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro Ser Thr  
 50 55 60  
 Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro Val Ala Pro Pro Pro Pro  
 65 70 75 80  
 Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro Gly Asp Pro Asn Ala Ala  
 85 90 95  
 Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro Pro Pro Val Ile Ala Pro  
 100 105 110  
 Asn Ala Pro Gln Pro Val Arg Ile Asp Asn Pro Val Gly Gly Phe Ser  
 115 120 125  
 Phe Ala Leu Pro Ala Gly Trp Val Glu Ser Asp Ala Ala His Phe Asp  
 130 135 140  
 Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr Gly Asp Pro Pro Phe Pro  
 145 150 155 160  
 Gly Gln Pro Pro Pro Val Ala Asn Asp Thr Arg Ile Val Leu Gly Arg  
 165 170 175  
 Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu Ala Thr Asp Ser Lys Ala  
 180 185 190  
 Ala Ala Arg Leu Gly Ser Asp Met Gly Glu Phe Tyr Met Pro Tyr Pro  
 195 200 205  
 Gly Thr Arg Ile Asn Gln Glu Thr Val Ser Leu Asp Ala Asn Gly Val  
 210 215 220  
 Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys  
 225 230 235 240  
 Pro Asn Gly Gln Ile Trp Thr Gly Val Ile Gly Ser Pro Ala Ala Asn  
 245 250 255  
 Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp Phe Val Val Trp Leu Gly  
 260 265 270  
 Thr Ala Asn Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Gln  
 275 280 285  
 Ser Ile Arg Pro Leu Val Ala Pro Pro Pro Ala Pro Ala Pro Ala Pro  
 290 295 300  
 Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro Thr  
 305 310 315 320  
 Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala  
 325 330

## (2) INFORMATION FOR SEQ ID NO:54:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Asn Tyr Gly Gln Val

1	5	10	15
Val	Ala	Ala	Leu
	20		

## (2) INFORMATION FOR SEQ ID NO:55:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ala	Val	Glu	Ser	Gly	Met	Leu	Ala	Leu	Gly	Thr	Pro	Ala	Pro	Ser
1				5					10					15

## (2) INFORMATION FOR SEQ ID NO:56:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 13 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ala	Ala	Met	Lys	Pro	Arg	Thr	Gly	Asp	Gly	Pro	Leu	Glu	Ala	Ala	Lys
1				5					10						15
Glu	Gly	Arg													

## (2) INFORMATION FOR SEQ ID NO:57:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Tyr	Tyr	Trp	Cys	Pro	Gly	Gln	Pro	Phe	Asp	Pro	Ala	Trp	Gly	Pro
1				5					10					15

## (2) INFORMATION FOR SEQ ID NO:58:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Asp Ile Gly Ser Glu Ser Thr Glu Asp Gln Gln Xaa Ala Val  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:59:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:60:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 17 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Ala Ala Ala Ala Pro Pro  
 1 5 10 15  
 Ala

(2) INFORMATION FOR SEQ ID NO:61:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly  
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:62:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Gln Thr Ser  
 1 5 10 15

Leu Leu Asn Asn Leu Ala Asp Pro Asp Val Ser Phe Ala Asp  
 20 25 30

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Thr Gly Ser Leu Asn Gln Thr His Asn Arg Arg Ala Asn Glu Arg Lys  
 1 5 10 15  
 Asn Thr Thr Met Lys Met Val Lys Ser Ile Ala Ala Gly Leu Thr Ala  
 20 25 30  
 Ala Ala Ala Ile Gly Ala Ala Ala Gly Val Thr Ser Ile Met Ala  
 35 40 45  
 Gly Gly Pro Val Val Tyr Gln Met Gln Pro Val Val Phe Gly Ala Pro  
 50 55 60  
 Leu Pro Leu Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln  
 65 70 75 80  
 Leu Thr Ser Leu Leu Asn Ser Leu Ala Asp Pro Asn Val Ser Phe Ala  
 85 90 95  
 Asn Lys Gly Ser Leu Val Glu Gly Gly Ile Gly Gly Thr Glu Ala Arg  
 100 105 110  
 Ile Ala Asp His Lys Leu Lys Lys Ala Ala Glu His Gly Asp Leu Pro  
 115 120 125  
 Leu Ser Phe Ser Val Thr Asn Ile Gln Pro Ala Ala Ala Gly Ser Ala  
 130 135 140  
 Thr Ala Asp Val Ser Val Ser Gly Pro Lys Leu Ser Ser Pro Val Thr  
 145 150 155 160  
 Gln Asn Val Thr Phe Val Asn Gln Gly Gly Trp Met Leu Ser Arg Ala  
 165 170 175  
 Ser Ala Met Glu Leu Leu Gln Ala Ala Gly Xaa  
 180 185

## (2) INFORMATION FOR SEQ ID NO:64:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Asp Glu Val Thr Val Glu Thr Thr Ser Val Phe Arg Ala Asp Phe Leu  
 1 5 10 15  
 Ser Glu Leu Asp Ala Pro Ala Gln Ala Gly Thr Glu Ser Ala Val Ser  
 20 25 30  
 Gly Val Glu Gly Leu Pro Pro Gly Ser Ala Leu Leu Val Val Lys Arg  
 35 40 45  
 Gly Pro Asn Ala Gly Ser Arg Phe Leu Leu Asp Gln Ala Ile Thr Ser

```

      50              55              60
Ala Gly Arg His Pro Asp Ser Asp Ile Phe Leu Asp Asp Val Thr Val
65              70              75              80
Ser Arg Arg His Ala Glu Phe Arg Leu Glu Asn Asn Glu Phe Asn Val
      85              90              95
Val Asp Val Gly Ser Leu Asn Gly Thr Tyr Val Asn Arg Glu Pro Val
      100              105              110
Asp Ser Ala Val Leu Ala Asn Gly Asp Glu Val Gln Ile Gly Lys Leu
      115              120              125
Arg Leu Val Phe Leu Thr Gly Pro Lys Gln Gly Glu Asp Asp Gly Ser
      130              135              140
Thr Gly Gly Pro
145

```

## (2) INFORMATION FOR SEQ ID NO:65:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```

Thr Ser Asn Arg Pro Ala Arg Arg Gly Arg Arg Ala Pro Arg Asp Thr
1      5      10      15
Gly Pro Asp Arg Ser Ala Ser Leu Ser Leu Val Arg His Arg Arg Gln
      20      25      30
Gln Arg Asp Ala Leu Cys Leu Ser Thr Gln Ile Ser Arg Gln Ser
      35      40      45
Asn Leu Pro Pro Ala Ala Gly Gly Ala Ala Asn Tyr Ser Arg Arg Asn
      50      55      60
Phe Asp Val Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu
      65      70      75      80
Leu Cys Cys Ser Gly Val Ala Thr Ala Ala Pro Lys Thr Tyr Cys Glu
      85      90      95
Glu Leu Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Gln Met Ser
      100      105      110
Asp Pro Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp
      115      120      125
Gln Lys Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu
      130      135      140
Ser Ala Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn
      145      150      155      160
Ile Thr Ser Ala Thr Tyr Gln Ser Ala Ile Pro Pro Arg Gly Thr Gln
      165      170      175
Ala Val Val Leu Xaa Val Tyr His Asn Ala Gly Gly Thr His Pro Thr
      180      185      190
Thr Thr Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile
      195      200      205
Thr Tyr Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val Val
      210      215      220
Phe Pro Ile Val Ala Arg
225      230

```

## (2) INFORMATION FOR SEQ ID NO:66:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

```

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1      5      10      15
Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20      25      30
Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35      40      45
Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50      55      60
Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65      70      75      80
Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85      90      95
Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
100     105     110
Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115     120     125
Gly Pro Pro Ala
130

```

## (2) INFORMATION FOR SEQ ID NO:67:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

```

Val Pro Leu Arg Ser Pro Ser Met Ser Pro Ser Lys Cys Leu Ala Ala
1      5      10      15
Ala Gln Arg Asn Pro Val Ile Arg Arg Arg Arg Leu Ser Asn Pro Pro
20      25      30
Pro Arg Lys Tyr Arg Ser Met Pro Ser Pro Ala Thr Ala Ser Ala Gly
35      40      45
Met Ala Arg Val Arg Arg Arg Ala Ile Trp Arg Gly Pro Ala Thr Xaa
50      55      60
Ser Ala Gly Met Ala Arg Val Arg Arg Trp Xaa Val Met Pro Xaa Val
65      70      75      80
Ile Gln Ser Thr Xaa Ile Arg Xaa Xaa Gly Pro Phe Asp Asn Arg Gly
85      90      95
Ser Glu Arg Lys
100

```

## (2) INFORMATION FOR SEQ ID NO:68:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```

Met Thr Asp Asp Ile Leu Leu Ile Asp Thr Asp Glu Arg Val Arg Thr
1           5           10           15
Leu Thr Leu Asn Arg Pro Gln Ser Arg Asn Ala Leu Ser Ala Ala Leu
20           25           30
Arg Asp Arg Phe Phe Ala Xaa Leu Xaa Asp Ala Glu Xaa Asp Asp Asp
35           40           45
Ile Asp Val Val Ile Leu Thr Gly Ala Asp Pro Val Phe Cys Ala Gly
50           55           60
Leu Asp Leu Lys Val Ala Gly Arg Ala Asp Arg Ala Ala Gly His Leu
65           70           75           80
Thr Ala Val Gly Gly His Asp Gln Ala Gly Asp Arg Arg Asp Gln Arg
85           90           95
Arg Arg Gly His Arg Arg Ala Arg Thr Gly Ala Val Leu Arg His Pro
100          105          110
Asp Arg Leu Arg Ala Arg Pro Leu Arg Arg His Pro Arg Pro Gly Gly
115          120          125
Ala Ala Ala His Leu Gly Thr Gln Cys Val Leu Ala Ala Lys Gly Arg
130          135          140
His Arg Xaa Gly Pro Val Asp Glu Pro Asp Arg Arg Leu Pro Val Arg
145          150          155          160
Asp Arg Arg

```

## (2) INFORMATION FOR SEQ ID NO:69:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```

Met Lys Phe Val Asn His Ile Glu Pro Val Ala Pro Arg Arg Ala Gly
1           5           10           15
Gly Ala Val Ala Glu Val Tyr Ala Glu Ala Arg Arg Glu Phe Gly Arg
20           25           30
Leu Pro Glu Pro Leu Ala Met Leu Ser Pro Asp Glu Gly Leu Leu Thr
35           40           45
Ala Gly Trp Ala Thr Leu Arg Glu Thr Leu Leu Val Gly Gln Val Pro
50           55           60
Arg Gly Arg Lys Glu Ala Val Ala Ala Ala Val Ala Ala Ser Leu Arg
65           70           75           80
Cys Pro Trp Cys Val Asp Ala His Thr Thr Met Leu Tyr Ala Ala Gly
85           90           95
Gln Thr Asp Thr Ala Ala Ala Ile Leu Ala Gly Thr Ala Pro Ala Ala
100          105          110

```

100

Gly Asp Pro Asn Ala Pro Tyr Val Ala Trp Ala Ala Gly Thr Gly Thr  
 115 120 125  
 Pro Ala Gly Pro Pro Ala Pro Phe Gly Pro Asp Val Ala Ala Glu Tyr  
 130 135 140  
 Leu Gly Thr Ala Val Glu Phe His Phe Ile Ala Arg Leu Val Leu Val  
 145 150 155 160  
 Leu Leu Asp Glu Thr Phe Leu Pro Gly Gly Pro Arg Ala Gln Gln Leu  
 165 170 175  
 Met Arg Arg Ala Gly Gly Leu Val Phe Ala Arg Lys Val Arg Ala Glu  
 180 185 190  
 His Arg Pro Gly Arg Ser Thr Arg Arg Leu Glu Pro Arg Thr Leu Pro  
 195 200 205  
 Asp Asp Leu Ala Trp Ala Thr Pro Ser Glu Pro Ile Ala Thr Ala Phe  
 210 215 220  
 Ala Ala Leu Ser His His Leu Asp Thr Ala Pro His Leu Pro Pro Pro  
 225 230 235 240  
 Thr Arg Gln Val Val Arg Arg Val Val Gly Ser Trp His Gly Glu Pro  
 245 250 255  
 Met Pro Met Ser Ser Arg Trp Thr Asn Glu His Thr Ala Glu Leu Pro  
 260 265 270  
 Ala Asp Leu His Ala Pro Thr Arg Leu Ala Leu Leu Thr Gly Leu Ala  
 275 280 285  
 Pro His Gln Val Thr Asp Asp Val Ala Ala Ala Arg Ser Leu Leu  
 290 295 300  
 Asp Thr Asp Ala Ala Leu Val Gly Ala Leu Ala Trp Ala Ala Phe Thr  
 305 310 315 320  
 Ala Ala Arg Arg Ile Gly Thr Trp Ile Gly Ala Ala Ala Glu Gly Gln  
 325 330 335  
 Val Ser Arg Gln Asn Pro Thr Gly  
 340

## (2) INFORMATION FOR SEQ ID NO:70:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Asp Asp Pro Asp Met Pro Gly Thr Val Ala Lys Ala Val Ala Asp Ala  
 1 5 10 15  
 Leu Gly Arg Gly Ile Ala Pro Val Glu Asp Ile Gln Asp Cys Val Glu  
 20 25 30  
 Ala Arg Leu Gly Glu Ala Gly Leu Asp Asp Val Ala Arg Val Tyr Ile  
 35 40 45  
 Ile Tyr Arg Gln Arg Arg Ala Glu Leu Arg Thr Ala Lys Ala Leu Leu  
 50 55 60  
 Gly Val Arg Asp Glu Leu Lys Leu Ser Leu Ala Ala Val Thr Val Leu  
 65 70 75 80  
 Arg Glu Arg Tyr Leu Leu His Asp Glu Gln Gly Arg Pro Ala Glu Ser  
 85 90 95  
 Thr Gly Glu Leu Met Asp Arg Ser Ala Arg Cys Val Ala Ala Ala Glu  
 100 105 110

Asp Gln Tyr Glu Pro Gly Ser Ser Arg Arg Trp Ala Glu Arg Phe Ala  
 115 120 125  
 Thr Leu Leu Arg Asn Leu Glu Phe Leu Pro Asn Ser Pro Thr Leu Met  
 130 135 140  
 Asn Ser Gly Thr Asp Leu Gly Leu Leu Ala Gly Cys Phe Val Leu Pro  
 145 150 155 160  
 Ile Glu Asp Ser Leu Gln Ser Ile Phe Ala Thr Leu Gly Gln Ala Ala  
 165 170 175  
 Glu Leu Gln Arg Ala Gly Gly Gly Thr Gly Tyr Ala Phe Ser His Leu  
 180 185 190  
 Arg Pro Ala Gly Asp Arg Val Ala Ser Thr Gly Gly Thr Ala Ser Gly  
 195 200 205  
 Pro Val Ser Phe Leu Arg Leu Tyr Asp Ser Ala Ala Gly Val Val Ser  
 210 215 220  
 Met Gly Gly Arg Arg Arg Gly Ala Cys Met Ala Val Leu Asp Val Ser  
 225 230 235 240  
 His Pro Asp Ile Cys Asp Phe Val Thr Ala Lys Ala Glu Ser Pro Ser  
 245 250 255  
 Glu Leu Pro His Phe Asn Leu Ser Val Gly Val Thr Asp Ala Phe Leu  
 260 265 270  
 Arg Ala Val Glu Arg Asn Gly Leu His Arg Leu Val Asn Pro Arg Thr  
 275 280 285  
 Gly Lys Ile Val Ala Arg Met Pro Ala Ala Glu Leu Phe Asp Ala Ile  
 290 295 300  
 Cys Lys Ala Ala His Ala Gly Gly Asp Pro Gly Leu Val Phe Leu Asp  
 305 310 315 320  
 Thr Ile Asn Arg Ala Asn Pro Val Pro Gly Arg Gly Arg Ile Glu Ala  
 325 330 335  
 Thr Asn Pro Cys Gly Glu Val Pro Leu Leu Pro Tyr Glu Ser Cys Asn  
 340 345 350  
 Leu Gly Ser Ile Asn Leu Ala Arg Met Leu Ala Asp Gly Arg Val Asp  
 355 360 365  
 Trp Asp Arg Leu Glu Glu Val Ala Gly Val Ala Val Arg Phe Leu Asp  
 370 375 380  
 Asp Val Ile Asp Val Ser Arg Tyr Pro Phe Pro Glu Leu Gly Glu Ala  
 385 390 395 400  
 Ala Arg Ala Thr Arg Lys Ile Gly Leu Gly Val Met Gly Leu Ala Glu  
 405 410 415  
 Leu Leu Ala Ala Leu Gly Ile Pro Tyr Asp Ser Glu Glu Ala Val Arg  
 420 425 430  
 Leu Ala Thr Arg Leu Met Arg Arg Ile Gln Gln Ala Ala His Thr Ala  
 435 440 445  
 Ser Arg Arg Leu Ala Glu Glu Arg Gly Ala Phe Pro Ala Phe Thr Asp  
 450 455 460  
 Ser Arg Phe Ala Arg Ser Gly Pro Arg Arg Asn Ala Gln Val Thr Ser  
 465 470 475 480  
 Val Ala Pro Thr Gly  
 485

(2) INFORMATION FOR SEQ ID NO:71:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 267 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single

102

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```

Gly Val Ile Val Leu Asp Leu Glu Pro Arg Gly Pro Leu Pro Thr Glu
1      5      10      15
Ile Tyr Trp Arg Arg Gly Leu Ala Leu Gly Ile Ala Val Val Val
20      25      30
Val Gly Ile Ala Val Ala Ile Val Ile Ala Phe Val Asp Ser Ser Ala
35      40      45
Gly Ala Lys Pro Val Ser Ala Asp Lys Pro Ala Ser Ala Gln Ser His
50      55      60
Pro Gly Ser Pro Ala Pro Gln Ala Pro Gln Pro Ala Gly Glu Thr Glu
65      70      75      80
Gly Asn Ala Ala Ala Ala Pro Pro Gln Gly Gln Asn Pro Glu Thr Pro
85      90      95
Thr Pro Thr Ala Ala Val Gln Pro Pro Pro Val Leu Lys Glu Gly Asp
100     105     110
Asp Cys Pro Asp Ser Thr Leu Ala Val Lys Gly Leu Thr Asn Ala Pro
115     120     125
Gln Tyr Tyr Val Gly Asp Gln Pro Lys Phe Thr Met Val Val Thr Asn
130     135     140
Ile Gly Leu Val Ser Cys Lys Arg Asp Val Gly Ala Ala Val Leu Ala
145     150     155     160
Ala Tyr Val Tyr Ser Leu Asp Asn Lys Arg Leu Trp Ser Asn Leu Asp
165     170     175
Cys Ala Pro Ser Asn Glu Thr Leu Val Lys Thr Phe Ser Pro Gly Glu
180     185     190
Gln Val Thr Thr Ala Val Thr Trp Thr Gly Met Gly Ser Ala Pro Arg
195     200     205
Cys Pro Leu Pro Arg Pro Ala Ile Gly Pro Gly Thr Tyr Asn Leu Val
210     215     220
Val Glu Leu Gly Asn Leu Arg Ser Leu Pro Val Pro Phe Ile Leu Asn
225     230     235     240
Gln Pro Pro Pro Pro Pro Gly Pro Val Pro Ala Pro Gly Pro Ala Gln
245     250     255
Ala Pro Pro Pro Glu Ser Pro Ala Gln Gly Gly
260     265

```

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 97 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

```

Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly Val Gln Val
1      5      10      15
Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu Val Val Ala
20      25      30
Gly Gly Ala Ala Ala Asn Ala Gly Val Pro Lys Gly Val Val Thr
35      40      45

```

Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu Val Ala Ala  
 50 55 60  
 Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr Phe Gln Asp  
 65 70 75 80  
 Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly Lys Ala Glu  
 85 90 95  
 Gln

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 364 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Gly Ala Ala Val Ser Leu Leu Ala Ala Gly Thr Leu Val Leu Thr Ala  
 1 5 10 15  
 Cys Gly Gly Gly Thr Asn Ser Ser Ser Ser Gly Ala Gly Gly Thr Ser  
 20 25 30  
 Gly Ser Val His Cys Gly Gly Lys Lys Glu Leu His Ser Ser Gly Ser  
 35 40 45  
 Thr Ala Gln Glu Asn Ala Met Glu Gln Phe Val Tyr Ala Tyr Val Arg  
 50 55 60  
 Ser Cys Pro Gly Tyr Thr Leu Asp Tyr Asn Ala Asn Gly Ser Gly Ala  
 65 70 75 80  
 Gly Val Thr Gln Phe Leu Asn Asn Glu Thr Asp Phe Ala Gly Ser Asp  
 85 90 95  
 Val Pro Leu Asn Pro Ser Thr Gly Gln Pro Asp Arg Ser Ala Glu Arg  
 100 105 110  
 Cys Gly Ser Pro Ala Trp Asp Leu Pro Thr Val Phe Gly Pro Ile Ala  
 115 120 125  
 Ile Thr Tyr Asn Ile Lys Gly Val Ser Thr Leu Asn Leu Asp Gly Pro  
 130 135 140  
 Thr Thr Ala Lys Ile Phe Asn Gly Thr Ile Thr Val Trp Asn Asp Pro  
 145 150 155 160  
 Gln Ile Gln Ala Leu Asn Ser Gly Thr Asp Leu Pro Pro Thr Pro Ile  
 165 170 175  
 Ser Val Ile Phe Arg Ser Asp Lys Ser Gly Thr Ser Asp Asn Phe Gln  
 180 185 190  
 Lys Tyr Leu Asp Gly Val Ser Asn Gly Ala Trp Gly Lys Gly Ala Ser  
 195 200 205  
 Glu Thr Phe Ser Gly Gly Val Gly Val Gly Ala Ser Gly Asn Asn Gly  
 210 215 220  
 Thr Ser Ala Leu Leu Gln Thr Thr Asp Gly Ser Ile Thr Tyr Asn Glu  
 225 230 235 240  
 Trp Ser Phe Ala Val Gly Lys Gln Leu Asn Met Ala Gln Ile Ile Thr  
 245 250 255  
 Ser Ala Gly Pro Asp Pro Val Ala Ile Thr Thr Glu Ser Val Gly Lys  
 260 265 270  
 Thr Ile Ala Gly Ala Lys Ile Met Gly Gln Gly Asn Asp Leu Val Leu  
 275 280 285  
 Asp Thr Ser Ser Phe Tyr Arg Pro Thr Gln Pro Gly Ser Tyr Pro Ile

104

```

      290              295              300
Val Leu Ala Thr Tyr Glu Ile Val Cys Ser Lys Tyr Pro Asp Ala Thr
305              310              315              320
Thr Gly Thr Ala Val Arg Ala Phe Met Gln Ala Ala Ile Gly Pro Gly
      325              330              335
Gln Glu Gly Leu Asp Gln Tyr Gly Ser Ile Pro Leu Pro Lys Ser Phe
      340              345              350
Gln Ala Lys Leu Ala Ala Ala Val Asn Ala Ile Ser
      355              360

```

## (2) INFORMATION FOR SEQ ID NO:74:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 309 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```

Gln Ala Ala Ala Gly Arg Ala Val Arg Arg Thr Gly His Ala Glu Asp
1           5           10           15
Gln Thr His Gln Asp Arg Leu His His Gly Cys Arg Arg Ala Ala Val
      20           25           30
Val Val Arg Gln Asp Arg Ala Ser Val Ser Ala Thr Ser Ala Arg Pro
      35           40           45
Pro Arg Arg His Pro Ala Gln Gly His Arg Arg Arg Val Ala Pro Ser
      50           55           60
Gly Gly Arg Arg Arg Pro His Pro His His Val Gln Pro Asp Asp Arg
      65           70           75           80
Arg Asp Arg Pro Ala Leu Leu Asp Arg Thr Gln Pro Ala Glu His Pro
      85           90           95
Asp Pro His Arg Arg Gly Pro Ala Asp Pro Gly Arg Val Arg Gly Arg
      100          105          110
Gly Arg Leu Arg Arg Val Asp Asp Gly Arg Leu Gln Pro Asp Arg Asp
      115          120          125
Ala Asp His Gly Ala Pro Val Arg Gly Arg Gly Pro His Arg Gly Val
      130          135          140
Gln His Arg Gly Gly Pro Val Phe Val Arg Arg Val Pro Gly Val Arg
      145          150          155          160
Cys Ala His Arg Arg Gly His Arg Arg Val Ala Ala Pro Gly Gln Gly
      165          170          175
Asp Val Leu Arg Ala Gly Leu Arg Val Glu Arg Leu Arg Pro Val Ala
      180          185          190
Ala Val Glu Asn Leu His Arg Gly Ser Gln Arg Ala Asp Gly Arg Val
      195          200          205
Phe Arg Pro Ile Arg Arg Gly Ala Arg Leu Pro Ala Arg Arg Ser Arg
      210          215          220
Ala Gly Pro Gln Gly Arg Leu His Leu Asp Gly Ala Gly Pro Ser Pro
      225          230          235          240
Leu Pro Ala Arg Ala Gly Gln Gln Gln Pro Ser Ser Ala Gly Gly Arg
      245          250          255
Arg Ala Gly Gly Ala Glu Arg Ala Asp Pro Gly Gln Arg Gly Arg His
      260          265          270
His Gln Gly Gly His Asp Pro Gly Arg Gln Gly Ala Gln Arg Gly Thr

```

[illegible]

(2) INFORMATION FOR SEQ ID NO: 73:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 580 amino acids

11. ~~12~~ ~~13~~ ~~14~~ ~~15~~ ~~16~~ ~~17~~ ~~18~~ ~~19~~ ~~20~~ ~~21~~ ~~22~~ ~~23~~ ~~24~~ ~~25~~ ~~26~~ ~~27~~ ~~28~~ ~~29~~ ~~30~~ ~~31~~ ~~32~~ ~~33~~ ~~34~~ ~~35~~ ~~36~~ ~~37~~ ~~38~~ ~~39~~ ~~40~~ ~~41~~ ~~42~~ ~~43~~ ~~44~~ ~~45~~ ~~46~~ ~~47~~ ~~48~~ ~~49~~ ~~50~~ ~~51~~ ~~52~~ ~~53~~ ~~54~~ ~~55~~ ~~56~~ ~~57~~ ~~58~~ ~~59~~ ~~60~~ ~~61~~ ~~62~~ ~~63~~ ~~64~~ ~~65~~ ~~66~~ ~~67~~ ~~68~~ ~~69~~ ~~70~~ ~~71~~ ~~72~~ ~~73~~ ~~74~~ ~~75~~ ~~76~~ ~~77~~ ~~78~~ ~~79~~ ~~80~~ ~~81~~ ~~82~~ ~~83~~ ~~84~~ ~~85~~ ~~86~~ ~~87~~ ~~88~~ ~~89~~ ~~90~~ ~~91~~ ~~92~~ ~~93~~ ~~94~~ ~~95~~ ~~96~~ ~~97~~ ~~98~~ ~~99~~ ~~100~~ ~~101~~ ~~102~~ ~~103~~ ~~104~~ ~~105~~ ~~106~~ ~~107~~ ~~108~~ ~~109~~ ~~110~~ ~~111~~ ~~112~~ ~~113~~ ~~114~~ ~~115~~ ~~116~~ ~~117~~ ~~118~~ ~~119~~ ~~120~~ ~~121~~ ~~122~~ ~~123~~ ~~124~~ ~~125~~ ~~126~~ ~~127~~ ~~128~~ ~~129~~ ~~130~~ ~~131~~ ~~132~~ ~~133~~ ~~134~~ ~~135~~ ~~136~~ ~~137~~ ~~138~~ ~~139~~ ~~140~~ ~~141~~ ~~142~~ ~~143~~ ~~144~~ ~~145~~ ~~146~~ ~~147~~ ~~148~~ ~~149~~ ~~150~~ ~~151~~ ~~152~~ ~~153~~ ~~154~~ ~~155~~ ~~156~~ ~~157~~ ~~158~~ ~~159~~ ~~160~~ ~~161~~ ~~162~~ ~~163~~ ~~164~~ ~~165~~ ~~166~~ ~~167~~ ~~168~~ ~~169~~ ~~170~~ ~~171~~ ~~172~~ ~~173~~ ~~174~~ ~~175~~ ~~176~~ ~~177~~ ~~178~~ ~~179~~ ~~180~~ ~~181~~ ~~182~~ ~~183~~ ~~184~~ ~~185~~ ~~186~~ ~~187~~ ~~188~~ ~~189~~ ~~190~~ ~~191~~ ~~192~~ ~~193~~ ~~194~~ ~~195~~ ~~196~~ ~~197~~ ~~198~~ ~~199~~ ~~200~~ ~~201~~ ~~202~~ ~~203~~ ~~204~~ ~~205~~ ~~206~~ ~~207~~ ~~208~~ ~~209~~ ~~210~~ ~~211~~ ~~212~~ ~~213~~ ~~214~~ ~~215~~ ~~216~~ ~~217~~ ~~218~~ ~~219~~ ~~220~~ ~~221~~ ~~222~~ ~~223~~ ~~224~~ ~~225~~ ~~226~~ ~~227~~ ~~228~~ ~~229~~ ~~230~~ ~~231~~ ~~232~~ ~~233~~ ~~234~~ ~~235~~ ~~236~~ ~~237~~ ~~238~~ ~~239~~ ~~240~~ ~~241~~ ~~242~~ ~~243~~ ~~244~~ ~~245~~ ~~246~~ ~~247~~ ~~248~~ ~~249~~ ~~250~~ ~~251~~ ~~252~~ ~~253~~ ~~254~~ ~~255~~ ~~256~~ ~~257~~ ~~258~~ ~~259~~ ~~260~~ ~~261~~ ~~262~~ ~~263~~ ~~264~~ ~~265~~ ~~266~~ ~~267~~ ~~268~~ ~~269~~ ~~270~~ ~~271~~ ~~272~~ ~~273~~ ~~274~~ ~~275~~ ~~276~~ ~~277~~ ~~278~~ ~~279~~ ~~280~~ ~~281~~ ~~282~~ ~~283~~ ~~284~~ ~~285~~ ~~286~~ ~~287~~ ~~288~~ ~~289~~ ~~290~~ ~~291~~ ~~292~~ ~~293~~ ~~294~~ ~~295~~ ~~296~~ ~~297~~ ~~298~~ ~~299~~ ~~300~~ ~~301~~ ~~302~~ ~~303~~ ~~304~~ ~~305~~ ~~306~~ ~~307~~ ~~308~~ ~~309~~ ~~310~~ ~~311~~ ~~312~~ ~~313~~ ~~314~~ ~~315~~ ~~316~~ ~~317~~ ~~318~~ ~~319~~ ~~320~~ ~~321~~ ~~322~~ ~~323~~ ~~324~~ ~~325~~ ~~326~~ ~~327~~ ~~328~~ ~~329~~ ~~330~~ ~~331~~ ~~332~~ ~~333~~ ~~334~~ ~~335~~ ~~336~~ ~~337~~ ~~338~~ ~~339~~ ~~340~~ ~~341~~ ~~342~~ ~~343~~ ~~344~~ ~~345~~ ~~346~~ ~~347~~ ~~348~~ ~~349~~ ~~350~~ ~~351~~ ~~352~~ ~~353~~ ~~354~~ ~~355~~ ~~356~~ ~~357~~ ~~358~~ ~~359~~ ~~360~~ ~~361~~ ~~362~~ ~~363~~ ~~364~~ ~~365~~ ~~366~~ ~~367~~ ~~368~~ ~~369~~ ~~370~~ ~~371~~ ~~372~~ ~~373~~ ~~374~~ ~~375~~ ~~376~~ ~~377~~ ~~378~~ ~~379~~ ~~380~~ ~~381~~ ~~382~~ ~~383~~ ~~384~~ ~~385~~ ~~386~~ ~~387~~ ~~388~~ ~~389~~ ~~390~~ ~~391~~ ~~392~~ ~~393~~ ~~394~~ ~~395~~ ~~396~~ ~~397~~ ~~398~~ ~~399~~ ~~400~~ ~~401~~ ~~402~~ ~~403~~ ~~404~~ ~~405~~ ~~406~~ ~~407~~ ~~408~~ ~~409~~ ~~410~~ ~~411~~ ~~412~~ ~~413~~ ~~414~~ ~~415~~ ~~416~~ ~~417~~ ~~418~~ ~~419~~ ~~420~~ ~~421~~ ~~422~~ ~~423~~ ~~424~~ ~~425~~ ~~426~~ ~~427~~ ~~428~~ ~~429~~ ~~430~~ ~~431~~ ~~432~~ ~~433~~ ~~434~~ ~~435~~ ~~436~~ ~~437~~ ~~438~~ ~~439~~ ~~440~~ ~~441~~ ~~442~~ ~~443~~ ~~444~~ ~~445~~ ~~446~~ ~~447~~ ~~448~~ ~~449~~ ~~450~~ ~~451~~ ~~452~~ ~~453~~ ~~454~~ ~~455~~ ~~456~~ ~~457~~ ~~458~~ ~~459~~ ~~460~~ ~~461~~ ~~462~~ ~~463~~ ~~464~~ ~~465~~ ~~466~~ ~~467~~ ~~468~~ ~~469~~ ~~470~~ ~~471~~ ~~472~~ ~~473~~ ~~474~~ ~~475~~

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 5.

Ser	Ala	Val	Trp	Cys	Leu	Asn	Gly	Phe	Thr	Gly	Arg	His	Arg	His	Gly
1				5					10					15	
Arg	Cys	Arg	Val	Arg	Ala	Ser	Gly	Trp	Arg	Ser	Ser	Asn	Arg	Trp	Cys
		20						25					30		
Ser	Thr	Thr	Ala	Asp	Cys	Cys	Ala	Ser	Lys	Thr	Pro	Thr	Gln	Ala	Ala
		35					40					45			
Ser	Pro	Leu	Glu	Arg	Arg	Phe	Thr	Cys	Cys	Ser	Pro	Ala	Val	Gly	Cys
		50				55					60				
Arg	Phe	Arg	Ser	Phe	Pro	Val	Arg	Arg	Leu	Ala	Leu	Gly	Ala	Arg	Thr
65					70				75					80	
Ser	Arg	Thr	Leu	Gly	Val	Arg	Arg	Thr	Leu	Ser	Gln	Trp	Asn	Leu	Ser
			85					90					95		
Pro	Arg	Ala	Gln	Pro	Ser	Cys	Ala	Val	Thr	Val	Glu	Ser	His	Thr	His
		100					105						110		
Ala	Ser	Pro	Arg	Met	Ala	Lys	Leu	Ala	Arg	Val	Val	Gly	Leu	Val	Gln
		115				120						125			
Glu	Glu	Gln	Pro	Ser	Asp	Met	Thr	Asn	His	Pro	Arg	Tyr	Ser	Pro	Pro
	130				135					140					
Pro	Glu	Gln	Pro	Gly	Thr	Pro	Gly	Tyr	Ala	Gln	Gly	Gln	Gln	Gln	Thr
145					150				155						160
Tyr	Ser	Glu	Gln	Phe	Asp	Trp	Arg	Tyr	Pro	Pro	Ser	Pro	Pro	Pro	Gln
				165				170						175	
Pro	Thr	Gln	Tyr	Arg	Gln	Pro	Tyr	Glu	Ala	Leu	Gly	Gly	Thr	Arg	Pro
		180					185						190		
Gly	Leu	Ile	Pro	Gly	Val	Ile	Pro	Thr	Met	Thr	Pro	Pro	Pro	Gly	Met
	195					200					205				
Val	Arg	Gln	Arg	Pro	Arg	Ala	Gly	Met	Leu	Ala	Ile	Gly	Ala	Val	Thr
	210					215					220				
Ile	Ala	Val	Val	Ser	Ala	Gly	Ile	Gly	Gly	Ala	Ala	Ala	Ser	Leu	Val
225					230					235					240
Gly	Phe	Asn	Arg	Ala	Pro	Ala	Gly	Pro	Ser	Gly	Gly	Pro	Val	Ala	Ala
				245				250					255		
Ser	Ala	Ala	Pro	Ser	Ile	Pro	Ala	Ala	Asn	Met	Pro	Pro	Gly	Ser	Val
		260					265						270		
Gln	Gln	Val	Ala	Ala	Lys	Val	Val	Pro	Ser	Val	Val	Met	Leu	Glu	Thr
		275				280						285			
Asp	Leu	Gly	Arg	Gln	Ser	Glu	Glu	Gly	Ser	Gly	Ile	Ile	Leu	Ser	Ala
	290					295					300				
Glu	Gly	Leu	Ile	Leu	Thr	Asn	Asn	His	Val	Ile	Ala	Ala	Ala	Ala	Lys

```

305          310          315          320
Pro Pro Leu Gly Ser Pro Pro Pro Lys Thr Thr Val Thr Phe Ser Asp
          325          330          335
Gly Arg Thr Ala Pro Phe Thr Val Val Gly Ala Asp Pro Thr Ser Asp
          340          345          350
Ile Ala Val Val Arg Val Gln Gly Val Ser Gly Leu Thr Pro Ile Ser
          355          360          365
Leu Gly Ser Ser Ser Asp Leu Arg Val Gly Gln Pro Val Leu Ala Ile
          370          375          380
Gly Ser Pro Leu Gly Leu Gln Gly Thr Val Thr Thr Gly Ile Val Ser
          385          390          395          400
Ala Leu Asn Arg Pro Val Ser Thr Thr Gly Glu Ala Gly Asn Gln Asn
          405          410          415
Thr Val Leu Asp Ala Ile Gln Thr Asp Ala Ala Ile Asn Pro Gly Asn
          420          425          430
Ser Gly Gly Ala Leu Val Asn Met Asn Ala Gln Leu Val Gly Val Asn
          435          440          445
Ser Ala Ile Ala Thr Leu Gly Ala Asp Ser Ala Asp Ala Gln Ser Gly
          450          455          460
Ser Ile Gly Leu Gly Phe Ala Ile Pro Val Asp Gln Ala Lys Arg Ile
          465          470          475          480
Ala Asp Glu Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly
          485          490          495
Val Gln Val Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu
          500          505          510
Val Val Ala Gly Gly Ala Ala Ala Asn Ala Gly Val Pro Lys Gly Val
          515          520          525
Val Val Thr Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu
          530          535          540
Val Ala Ala Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr
          545          550          555          560
Phe Gln Asp Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly
          565          570          575
Lys Ala Glu Gln
          580

```

## (2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 233 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

Met Asn Asp Gly Lys Arg Ala Val Thr Ser Ala Val Leu Val Val Leu
1           5           10           15
Gly Ala Cys Leu Ala Leu Trp Leu Ser Gly Cys Ser Ser Pro Lys Pro
20           25           30
Asp Ala Glu Glu Gln Gly Val Pro Val Ser Pro Thr Ala Ser Asp Pro
35           40           45
Ala Leu Leu Ala Glu Ile Arg Gln Ser Leu Asp Ala Thr Lys Gly Leu
50           55           60
Thr Ser Val His Val Ala Val Arg Thr Thr Gly Lys Val Asp Ser Leu

```

65	70	75	80
Leu Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala Asn Pro Leu Ala			
85	90	95	
Ala Lys Gly Val Cys Thr Tyr Asn Asp Glu Gln Gly Val Pro Phe Arg			
100	105	110	
Val Gln Gly Asp Asn Ile Ser Val Lys Leu Phe Asp Asp Trp Ser Asn			
115	120	125	
Leu Gly Ser Ile Ser Glu Leu Ser Thr Ser Arg Val Leu Asp Pro Ala			
130	135	140	
Ala Gly Val Thr Gln Leu Leu Ser Gly Val Thr Asn Leu Gln Ala Gln			
145	150	155	160
Gly Thr Glu Val Ile Asp Gly Ile Ser Thr Thr Lys Ile Thr Gly Thr			
165	170	175	
Ile Pro Ala Ser Ser Val Lys Met Leu Asp Pro Gly Ala Lys Ser Ala			
180	185	190	
Arg Pro Ala Thr Val Trp Ile Ala Gln Asp Gly Ser His His Leu Val			
195	200	205	
Arg Ala Ser Ile Asp Leu Gly Ser Gly Ser Ile Gln Leu Thr Gln Ser			
210	215	220	
Lys Trp Asn Glu Pro Val Asn Val Asp			
225	230		

## (2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 66 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Val Ile Asp Ile Ile Gly Thr Ser Pro Thr Ser Trp Glu Gln Ala Ala		
1	5	10
Ala Glu Ala Val Gln Arg Ala Arg Asp Ser Val Asp Asp Ile Arg Val		
20	25	30
Ala Arg Val Ile Glu Gln Asp Met Ala Val Asp Ser Ala Gly Lys Ile		
35	40	45
Thr Tyr Arg Ile Lys Leu Glu Val Ser Phe Lys Met Arg Pro Ala Gln		
50	55	60
Pro Arg		
65		

## (2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 69 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Val Pro Pro Ala Pro Pro Leu Pro Pro Leu Pro Pro Ser Pro Ile Ser		
1	5	10
		15

Cys Ala Ser Pro Pro Ser Pro Pro Leu Pro Pro Ala Pro Pro Val Ala  
 20 25 30  
 Pro Gly Pro Pro Met Pro Pro Leu Asp Pro Trp Pro Pro Ala Pro Pro  
 35 40 45  
 Leu Pro Tyr Ser Thr Pro Pro Gly Ala Pro Leu Pro Pro Ser Pro Pro  
 50 55 60  
 Ser Pro Pro Leu Pro  
 65

## (2) INFORMATION FOR SEQ ID NO:79:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser  
 1 5 10 15  
 Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala  
 20 25 30  
 Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu  
 35 40 45  
 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val  
 50 55 60  
 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr  
 65 70 75 80  
 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val  
 85 90 95  
 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln  
 100 105 110  
 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala  
 115 120 125  
 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly  
 130 135 140  
 Gly Gly Val Ala Val Gly Gln Pro Val Val Ala Met Gly Asn Ser Gly  
 145 150 155 160  
 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu  
 165 170 175  
 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr  
 180 185 190  
 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser  
 195 200 205  
 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr  
 210 215 220  
 Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala  
 225 230 235 240  
 Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly  
 245 250 255  
 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu  
 260 265 270  
 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val  
 275 280 285

109

Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile  
 290 295 300  
 Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp  
 305 310 315 320  
 Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln  
 325 330 335  
 Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly  
 340 345 350  
 Pro Pro Ala  
 355

## (2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 205 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Ser Pro Lys Pro Asp Ala Glu Glu Gln Gly Val Pro Val Ser Pro Thr  
 1 5 10 15  
 Ala Ser Asp Pro Ala Leu Leu Ala Glu Ile Arg Gln Ser Leu Asp Ala  
 20 25 30  
 Thr Lys Gly Leu Thr Ser Val His Val Ala Val Arg Thr Thr Gly Lys  
 35 40 45  
 Val Asp Ser Leu Leu Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala  
 50 55 60  
 Asn Pro Leu Ala Ala Lys Gly Val Cys Thr Tyr Asn Asp Glu Gln Gly  
 65 70 75 80  
 Val Pro Phe Arg Val Gln Gly Asp Asn Ile Ser Val Lys Leu Phe Asp  
 85 90 95  
 Asp Trp Ser Asn Leu Gly Ser Ile Ser Glu Leu Ser Thr Ser Arg Val  
 100 105 110  
 Leu Asp Pro Ala Ala Gly Val Thr Gln Leu Leu Ser Gly Val Thr Asn  
 115 120 125  
 Leu Gln Ala Gln Gly Thr Glu Val Ile Asp Gly Ile Ser Thr Thr Lys  
 130 135 140  
 Ile Thr Gly Thr Ile Pro Ala Ser Ser Val Lys Met Leu Asp Pro Gly  
 145 150 155 160  
 Ala Lys Ser Ala Arg Pro Ala Thr Val Trp Ile Ala Gln Asp Gly Ser  
 165 170 175  
 His His Leu Val Arg Ala Ser Ile Asp Leu Gly Ser Gly Ser Ile Gln  
 180 185 190  
 Leu Thr Gln Ser Lys Trp Asn Glu Pro Val Asn Val Asp  
 195 200 205

## (2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 286 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Gly Asp Ser Phe Trp Ala Ala Ala Asp Gln Met Ala Arg Gly Phe Val  
 1 5 10 15  
 Leu Gly Ala Thr Ala Gly Arg Thr Thr Leu Thr Gly Glu Gly Leu Gln  
 20 25 30  
 His Ala Asp Gly His Ser Leu Leu Leu Asp Ala Thr Asn Pro Ala Val  
 35 40 45  
 Val Ala Tyr Asp Pro Ala Phe Ala Tyr Glu Ile Gly Tyr Ile Xaa Glu  
 50 55 60  
 Ser Gly Leu Ala Arg Met Cys Gly Glu Asn Pro Glu Asn Ile Phe Phe  
 65 70 75 80  
 Tyr Ile Thr Val Tyr Asn Glu Pro Tyr Val Gln Pro Pro Glu Pro Glu  
 85 90 95  
 Asn Phe Asp Pro Glu Gly Val Leu Gly Gly Ile Tyr Arg Tyr His Ala  
 100 105 110  
 Ala Thr Glu Gln Arg Thr Asn Lys Xaa Gln Ile Leu Ala Ser Gly Val  
 115 120 125  
 Ala Met Pro Ala Ala Leu Arg Ala Ala Gln Met Leu Ala Ala Glu Trp  
 130 135 140  
 Asp Val Ala Ala Asp Val Trp Ser Val Thr Ser Trp Gly Glu Leu Asn  
 145 150 155 160  
 Arg Asp Gly Val Val Ile Glu Thr Glu Lys Leu Arg His Pro Asp Arg  
 165 170 175  
 Pro Ala Gly Val Pro Tyr Val Thr Arg Ala Leu Glu Asn Ala Arg Gly  
 180 185 190  
 Pro Val Ile Ala Val Ser Asp Trp Met Arg Ala Val Pro Glu Gln Ile  
 195 200 205  
 Arg Pro Trp Val Pro Gly Thr Tyr Leu Thr Leu Gly Thr Asp Gly Phe  
 210 215 220  
 Gly Phe Ser Asp Thr Arg Pro Ala Gly Arg Arg Tyr Phe Asn Thr Asp  
 225 230 235 240  
 Ala Glu Ser Gln Val Gly Arg Gly Phe Gly Arg Gly Trp Pro Gly Arg  
 245 250 255  
 Arg Val Asn Ile Asp Pro Phe Gly Ala Gly Arg Gly Pro Pro Ala Glu  
 260 265 270  
 Leu Pro Gly Phe Asp Glu Gly Gly Gly Leu Arg Pro Xaa Lys  
 275 280 285

## (2) INFORMATION FOR SEQ ID NO:82:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 173 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Thr Lys Phe His Ala Leu Met Gln Gln Gln Ile His Asn Glu Phe Thr  
 1 5 10 15  
 Ala Ala Gln Gln Tyr Val Ala Ile Ala Val Tyr Phe Asp Ser Glu Asp  
 20 25 30  
 Leu Pro Gln Leu Ala Lys His Phe Tyr Ser Gln Ala Val Glu Glu Arg

111

```

      35              40              45
Asn His Ala Met Met Leu Val Gln His Leu Leu Asp Arg Asp Leu Arg
 55              60
Val Glu Ile Pro Gly Val Asp Thr Val Arg Asn Gln Phe Asp Arg Pro
 65              70              75              80
Arg Glu Ala Leu Ala Leu Ala Leu Asp Gln Glu Arg Thr Val Thr Asp
      85              90              95
Gln Val Gly Arg Leu Thr Ala Val Ala Arg Asp Glu Gly Asp Phe Leu
      100              105              110
Gly Glu Gln Phe Met Gln Trp Phe Leu Gln Glu Gln Ile Glu Glu Val
      115              120              125
Ala Leu Met Ala Thr Leu Val Arg Val Ala Asp Arg Ala Gly Ala Asn
      130              135              140
Leu Phe Glu Leu Glu Asn Phe Val Ala Arg Glu Val Asp Val Ala Pro
      145              150              155              160
Ala Ala Ser Gly Ala Pro His Ala Ala Gly Gly Arg Leu
      165              170

```

## (2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 107 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

```

Arg Ala Asp Glu Arg Lys Asn Thr Thr Met Lys Met Val Lys Ser Ile
 1              5              10              15
Ala Ala Gly Leu Thr Ala Ala Ala Ala Ile Gly Ala Ala Ala Gly
      20              25              30
Val Thr Ser Ile Met Ala Gly Gly Pro Val Val Tyr Gln Met Gln Pro
      35              40              45
Val Val Phe Gly Ala Pro Leu Pro Leu Asp Pro Xaa Ser Ala Pro Xaa
      50              55              60
Val Pro Thr Ala Ala Gln Trp Thr Xaa Leu Leu Asn Xaa Leu Xaa Asp
      65              70              75              80
Pro Asn Val Ser Phe Xaa Asn Lys Gly Ser Leu Val Glu Gly Gly Ile
      85              90              95
Gly Gly Xaa Glu Gly Xaa Xaa Arg Arg Xaa Gln
      100              105

```

## (2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 125 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

```

Val Leu Ser Val Pro Val Gly Asp Gly Phe Trp Xaa Arg Val Val Asn
 1              5              10              15
Pro Leu Gly Gln Pro Ile Asp Gly Arg Gly Asp Val Asp Ser Asp Thr

```

112

```

      20      25      30
Arg Arg Ala Leu Glu Leu Gln Ala Pro Ser Val Val Xaa Arg Gln Gly
      35      40      45
Val Lys Glu Pro Leu Xaa Thr Gly Ile Lys Ala Ile Asp Ala Met Thr
      50      55      60
Pro Ile Gly Arg Gly Gln Arg Gln Leu Ile Ile Gly Asp Arg Lys Thr
      65      70      75      80
Gly Lys Asn Arg Arg Leu Cys Arg Thr Pro Ser Ser Asn Gln Arg Glu
      85      90      95
Glu Leu Gly Val Arg Trp Ile Pro Arg Ser Arg Cys Ala Cys Val Tyr
      100      105      110
Val Gly His Arg Ala Arg Arg Gly Thr Tyr His Arg Arg
      115      120      125

```

## (2) INFORMATION FOR SEQ ID NO:85:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

```

Cys Asp Ala Val Met Gly Phe Leu Gly Gly Ala Gly Pro Leu Ala Val
1      5      10      15
Val Asp Gln Gln Leu Val Thr Arg Val Pro Gln Gly Trp Ser Phe Ala
      20      25      30
Gln Ala Ala Ala Val Pro Val Val Phe Leu Thr Ala Trp Tyr Gly Leu
      35      40      45
Ala Asp Leu Ala Glu Ile Lys Ala Gly Glu Ser Val Leu Ile His Ala
      50      55      60
Gly Thr Gly Gly Val Gly Met Ala Ala Val Gln Leu Ala Arg Gln Trp
      65      70      75      80
Gly Val Glu Val Phe Val Thr Ala Ser Arg Gly Lys Trp Asp Thr Leu
      85      90      95
Arg Ala Xaa Xaa Phe Asp Asp Xaa Pro Tyr Arg Xaa Phe Pro His Xaa
      100      105      110
Arg Ser Ser Xaa Gly
      115

```

## (2) INFORMATION FOR SEQ ID NO:86:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```

Met Tyr Arg Phe Ala Cys Arg Thr Leu Met Leu Ala Ala Cys Ile Leu
1      5      10      15
Ala Thr Gly Val Ala Gly Leu Gly Val Gly Ala Gln Ser Ala Ala Gln
      20      25      30

```

```

Thr Ala Pro Val Pro Asp Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp
    35                                40                                45
Pro Ala Trp Gly Pro Asn Trp Asp Pro Tyr Thr Cys His Asp Asp Phe
    50                                55                                60
His Arg Asp Ser Asp Gly Pro Asp His Ser Arg Asp Tyr Pro Gly Pro
    65                                70                                75                                80
Ile Leu Glu Gly Pro Val Leu Asp Asp Pro Gly Ala Ala Pro Pro Pro
    85                                90                                95
Pro Ala Ala Gly Gly Gly Ala
    100

```

## (2) INFORMATION FOR SEQ ID NO:87:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```

Val Gln Cys Arg Val Trp Leu Glu Ile Gln Trp Arg Gly Met Leu Gly
1      5      10      15
Ala Asp Gln Ala Arg Ala Gly Gly Pro Ala Arg Ile Trp Arg Glu His
    20      25      30
Ser Met Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala
    35      40      45
Thr Lys Glu Gly Arg Gly Ile Val Met Arg Val Pro Leu Glu Gly Gly
    50      55      60
Gly Arg Leu Val Val Glu Leu Thr Pro Asp Glu Ala Ala Ala Leu Gly
    65      70      75      80
Asp Glu Leu Lys Gly Val Thr Ser
    85

```

## (2) INFORMATION FOR SEQ ID NO:88:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```

Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile
1      5      10      15
Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly
    20      25      30
Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala
    35      40      45
Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys Gln Lys Gln Glu Leu
    50      55      60
Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg
    65      70      75      80
Ala Asp Glu Glu Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe
    85

```

85

90

95

## (2) INFORMATION FOR SEQ ID NO:89:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

```

Met Thr Gln Ser Gln Thr Val Thr Val Asp Gln Gln Glu Ile Leu Asn
 1           5           10           15
Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val
          20           25           30
Pro Ile Thr Pro Cys Glu Leu Thr Xaa Xaa Lys Asn Ala Ala Gln Gln
          35           40           45
Xaa Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala
          50           55           60
Lys Glu Arg Gln Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Xaa
          65           70           75           80
Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly
          85           90           95
Glu Gly Thr Val Gln Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser
          100          105          110
Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro
          115          120          125
Asn Phe Met Asp Leu Lys Gln Ala Ala Arg Lys Leu Glu Thr Gly Asp
          130          135          140
Gln Gly Ala Ser Leu Ala His Xaa Gly Asp Gly Trp Asn Thr Xaa Thr
          145          150          155          160
Leu Thr Leu Gln Gly Asp
          166

```

## (2) INFORMATION FOR SEQ ID NO:90:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

```

Arg Ala Glu Arg Met
 1           5

```

## (2) INFORMATION FOR SEQ ID NO:91:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
1      5      10      15
Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
20      25      30
Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
35      40      45
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
50      55      60
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
65      70      75      80
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
85      90      95
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
100      105      110
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
115      120      125
Asn Asn Val Pro Gln Ala Leu Lys Gln Leu Ala Gln Pro Thr Gln Gly
130      135      140
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
145      150      155      160
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
165      170      175
Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
180      185      190
Leu Lys Gly Phe Ala Pro Ala Ala Ala Gln Ala Val Gln Thr Ala
195      200      205
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
210      215      220
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
225      230      235      240
Ser Val Arg Tyr Gly His Arg Asp Gly Gly Lys Tyr Ala Lys Ser Gly
245      250      255
Arg Arg Asn Gly Pro Ala
260

```

## (xii) INFORMATION FOR SEQ ID NO:92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 303 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```

Met Thr Tyr Ser Pro Gly Asn Pro Gly Tyr Pro Gln Ala Gln Pro Ala
1      5      10      15
Gly Ser Tyr Gly Gly Val Thr Pro Ser Phe Ala His Ala Asp Glu Gly
20      25      30
Ala Ser Lys Leu Pro Met Tyr Leu Asn Ile Ala Val Ala Val Leu Gly
35      40      45
Leu Ala Ala Tyr Phe Ala Ser Phe Gly Pro Met Phe Thr Leu Ser Thr

```

116

```

      50      55      60
Glu Leu Gly Gly Gly Asp Gly Ala Val Ser Gly Asp Thr Gly Leu Pro
65      70      75      80
Val Gly Val Ala Leu Leu Ala Ala Leu Leu Ala Gly Val Val Leu Val
      85      90      95
Pro Lys Ala Lys Ser His Val Thr Val Val Ala Val Leu Gly Val Leu
      100      105      110
Gly Val Phe Leu Met Val Ser Ala Thr Phe Asn Lys Pro Ser Ala Tyr
      115      120      125
Ser Thr Gly Trp Ala Leu Trp Val Val Leu Ala Phe Ile Val Phe Gln
      130      135      140
Ala Val Ala Ala Val Leu Ala Leu Leu Val Glu Thr Gly Ala Ile Thr
      145      150      155      160
Ala Pro Ala Pro Arg Pro Lys Phe Asp Pro Tyr Gly Gln Tyr Gly Arg
      165      170      175
Tyr Gly Gln Tyr Gly Gln Tyr Gly Val Gln Pro Gly Gly Tyr Tyr Gly
      180      185      190
Gln Gln Gly Ala Gln Gln Ala Ala Gly Leu Gln Ser Pro Gly Pro Gln
      195      200      205
Gln Ser Pro Gln Pro Pro Gly Tyr Gly Ser Gln Tyr Gly Gly Tyr Ser
      210      215      220
Ser Ser Pro Ser Gln Ser Gly Ser Gly Tyr Thr Ala Gln Pro Pro Ala
      225      230      235      240
Gln Pro Pro Ala Gln Ser Gly Ser Gln Gln Ser His Gln Gly Pro Ser
      245      250      255
Thr Pro Pro Thr Gly Phe Pro Ser Phe Ser Pro Pro Pro Pro Val Ser
      260      265      270
Ala Gly Thr Gly Ser Gln Ala Gly Ser Ala Pro Val Asn Tyr Ser Asn
      275      280      285
Pro Ser Gly Gly Glu Gln Ser Ser Ser Pro Gly Gly Ala Pro Val
      290      295      300

```

## (2) INFORMATION FOR SEQ ID NO:93:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```

Gly Cys Gly Glu Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn
1      5      10      15
Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile
      20      25

```

## (2) INFORMATION FOR SEQ ID NO:94:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```

Asp Gln Val Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly
1           5           10           15

```

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```

Gly Cys Gly Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala
1           5           10           15
Ala Gly Thr Ala Ala Gln Ala Ala Val Val Arg
20           25

```

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

```

Gly Cys Gly Gly Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu
1           5           10           15
Ala Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu
20           25

```

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```

Gly Cys Gly Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu Ile Ser Thr
1           5           10           15
Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg
20           25

```

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 amino acids

- (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Gly	Cys	Gly	Ile	Arg	Gln	Ala	Gly	Val	Gln	Tyr	Ser	Arg	Ala	Asp	Glu
1				5				10						15	
Glu	Gln	Gln	Gln	Ala	Leu	Ser	Ser	Gln	Met	Gly	Phe				
			20					25							

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 507 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

ATGAAGATGG	TGAAATCCAT	CGCCGCAGGT	CTGACCCGCC	CGGCTGCAAT	CGGCCCCGCT	60
GGGCCCCGTC	TGACTTCGAT	CATGGCTGCC	GGCCCCGGTC	TATACCAGAT	GCAGCCCGTC	120
GTCTTCGGCG	CGCCTCTGCC	GTTGGAGCCG	GCTTCCGCCG	CTGACCTGCC	GACCCCCGCT	180
CAGTTGACCA	GCTTCTCAA	CAGCTTCGCC	GATCCCAACC	TGTCCTTTCG	GAACAAGGTC	240
AGTCTGCTCG	AGGCGGGCAT	CGGGGCGACC	GAGGCGCGCA	TGCCCCACCA	CAAGCTGAAG	300
AAGGCCCCCG	AGCAGCGGGA	TCTGCTCGTC	TCTTTCAGCG	TGACGAACAT	CGAGCCCGCG	360
GGGCCCCGTT	CGGCACCCGC	CGACCTTTCG	GTCTCGGGTC	CGAAGCTCTC	GTGCCCCGTC	420
ACGAGAGACG	TCAGCTTCGT	GAATCAAGGC	GGCTGGATGC	TGTCACGCGC	ATCGCCGATG	480
GAGTTGCTGC	AGGCCGACGG	GAAGTGA				507

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 158 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Met	Lys	Met	Val	Lys	Ser	Ile	Ala	Ala	Gly	Leu	Thr	Ala	Ala	Ala	Ala
1			5						10					15	
Ile	Gly	Ala	Ala	Ala	Ala	Gly	Val	Thr	Ser	Ile	Met	Ala	Gly	Gly	Pro
			20					25				30			
Val	Val	Tyr	Gln	Met	Gln	Pro	Val	Val	Phe	Gly	Ala	Pro	Leu	Pro	Leu
		35				40			45						
Asp	Pro	Ala	Ser	Ala	Pro	Asp	Val	Pro	Thr	Ala	Ala	Gln	Leu	Thr	Ser
	50			55				60							
Leu	Leu	Asn	Ser	Leu	Ala	Asp	Pro	Asn	Val	Ser	Phe	Ala	Asn	Lys	Gly
	65			70				75				80			
Ser	Leu	Val	Glu	Gly	Gly	Ile	Gly	Gly	Thr	Glu	Ala	Arg	Ile	Ala	Asp
			85				90					95			
His	Lys	Leu	Lys	Lys	Ala	Ala	Glu	His	Gly	Asp	Leu	Pro	Leu	Ser	Phe

119

	100		105		110
Ser Val Thr Asn Ile Gln Pro Ala Ala Ala Gly Ser Ala Thr Ala Asp					
115		120		125	
Val Ser Val Ser Gly Pro Lys Leu Ser Ser Pro Val Thr Gln Asn Val					
130		135		140	
Thr Phe Val Asn Gln Gly Gly Trp Met Leu Ser Arg Ala Ser Ala Met					
145		150		155	160
Glu Leu Leu Gln Ala Ala Gly Asn					
	165				

## (2) INFORMATION FOR SEQ ID NO:101:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CGTGGCAATG TCGTTGACCG TCGGGGCGCG GGTGCGCTCC GCAGATCCCG TGGACGCGGT	60
CATTAAACCC ACCTGCAATY ACGGGCAGGT AGTAGCTGCC CTCACGCGCA CGGATCGGGG	120
GGCTGCCGCA CATTCAACG CTTACCCGGT GGCCAGTCC TATTTCGCGA ATTTCTCTGC	180
CGCACCGCCA CTTACGCGCG CTGCCATGGC CGCGCAATTG CAAGCTGTGC CCGGGGCGGC	240
ACAGTACATC GGCCTTGTCT AGTCGTTTGC CGGCTCTTGC AACAACTATT AAGCCCATGC	300
GGGCGCCATC CCGCGACCGG GCATCCTTGC CCGGCTTAGG CCAGATTGCC CCGCTCTTCA	360
ACGGCGCGCA TCCCGCGACC CGGCATCGTC GCGGGGCGTA GGCCAGATTG CCGGCTCTCT	420
CAACGGGCGG CATCTCTTGC CGAATTCCTG CAGCCCGCGG GATCCACTAG TTCTAGAGCG	480
GGCGCCACCG CGCTCGAGCT	500

## (2) INFORMATION FOR SEQ ID NO:102:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 96 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Val Ala Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro	
1	15
Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala	
20	30
Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser	
35	45
Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro	
50	60
Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala	
65	80
Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr	
85	95

## (2) INFORMATION FOR SEQ ID NO:103:

120

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```

ATGACAGAGC AGCAGTGGAA TTTCGCGGCT ATCGAGGCGG CGGCAAGCGC AATCCAGGGA      60
AATGTACGCT CCATTGATTC CTCCTTGAC GAGGGGAAGC AGTCCCTGAC CAGCTCGCA      120
GCGGCTTGGG GCGGTAGCGG TTGGGAAGCG TACC                                154

```

## (2) INFORMATION FOR SEQ ID NO:104:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

```

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser
 1             5             10             15
Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly
      20             25             30
Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser
      35             40             45
Glu Ala Tyr
      50

```

## (2) INFORMATION FOR SEQ ID NO:105:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 282 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

```

CGGTCCGCGCA CTTCAGGTC ACTATGAAAG TCGGCTTCCG NCTGGAGGAT TCTGAACTT      60
TCAAGCGCGG CGGTAAGTC AGGTGCATCA TTAAGCGACT TTTCGAGAAC ATCTGACGC      120
GCTCGAAACE CGGCACAGCC GACGCTGGCT CCGGCGAGGC GCTGCTCCA AAATCTCTGA      180
GACAATTCGN CGGGGCGGCC TACAAGGAAG TCGGTGCTGA ATTGCGGNG TATCTGGTGG      240
ACCTGTGTGG TCTGAGGCC GACGAAGCGG TGCTGACGT CG                                282

```

## (2) INFORMATION FOR SEQ ID NO:106:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3058 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GATCGTACCC GTCCGAGTGC TCGGCCCGTT TGAGGATGGA GTGCACGTGT CTTTCGTGAT	60
GGCATAACCA GAGATGTTGG CGCGCGCGGC TGACACCGCTG CAGAGCATCG GTGCTACCAAC	120
TGTGGCTAGC AATGCCGCTG CGCGCGCGCC GAGGACTGGG GTGGTGGCCC CCGGTGCCGA	180
TGAGGTGTGG GCGCTGACTG CCGCGCACTT CCGCGCACAT GCGGCGATGT ATCAGTCCCT	240
GAGCGCTCGG GCTGCTGCGA TTCATGACCA GTTCGTGGCC ACCCTTGCCA GCAGCGGCCAG	300
CTCTATGCG GCCACTGAGG TCGCCAATGC GCGCGCGGCC AGCTAAGCCA GGACAGTCC	360
GCACGAGAAA CCACGAGAAA TAGGGACAGG TAATGGTGGG TTTCGGGGCG TTACCAACCG	420
AGATCAACTC CCGCAGGATG TACCGCGGCC CCGGTTCCGC CTCGCTGGTG GCGCGGCTT	480
AGATGTGGGA CAGCGTGCGG AGTGACCTGT TTTCGGGGCG GTCGGCGTTT CAGTCGTGG	540
TCTGGGGTCT GACCGTGGGG TCGTGGATAG GTTCGTCCGC GGTCTGATG GTGGCGGCGG	600
CTTCGCGGTA TGTGGCTGG ATGAGCGTCA CCGCGGGGCA GCGCGAGCTG ACCGCGGCCC	660
AGGTCCGGGT TGCTGCGCGG GCTTACGAGA CCGGCTATGG GTTGACGGTG CCGCGCGCGG	720
TGATCGCGGA GAACCTTGCT GAACGTATGA TTCTGATAGC GACCAACCTC TTGGGGCAAA	780
ACACCCCGGC GATCGCGGTC AAGGAGGCGG AATACGCGA GATGTGGGCC CAAGACGCCC	840
CGCGCATGTT TGCTACGCC CGCGCGACCG CGACCGCGAC GCGGACGTTG GTGCGCTTCG	900
AGGAGGCGCC GGAGATGACC AGCGCGGGTG GGCTCTCGA GCAGGCGCGC GCGGTCCAGG	960
AGGCTGCGGA CACCGCGCGG GCGAACCACT TGATGAACAA TTGCGCCAGG GCGCTGCAAC	1020
AGCTGGCCCA GCGCACGAGG GGCACGAGC CTTCTTCCAA GCTGGGTGGC CTGTGGAAGA	1080
CGGTCTCGCC GCATCGGTGG CCGATCAGCA ACATGGTCTC GATGGCCAAC AACCACTGT	1140
CGATGACCAA CTCGGGTGTG TCGATGACCA ACACCTTGAG CTCGATGTTG AAGGCTTTG	1200
CTCGCGCGGC GCGCGCGCAG GCGGTCCAAA CCGCGCGCCA AAACGGGCTC CCGGCGATGA	1260
GCTCGGTGGG CAGCTCGCTG GGTTCCTTGG GTCTGGGCGG TGGGTGGGCC GCCAAGTTGG	1320
GTGGGCGCGC CTCGGTGGGT TCGTGTTCGG TCGCGCAGGC CTGGGCGCGC GCCAAGCAGG	1380
CAGTCACCCC GCGCGCGCGG GCGCTGCGCG TGACGAGCCT GACGAGCGCC GCGGAAGAG	1440
GGCGCGCGCA GATGCTGGGC GCGCTGCGCG TGGGCGAGAT GCGCGCGCAG GCCGCTGGTG	1500
GGCTCAGTGG TGTGCTGCTT GTTCGCGCGC GACCTATGAT GATGCGCGAT TCTCGCGCGG	1560
CGGCTAGGA GAGGGGCGCG AACTGTCTCT TATTTGACCA GTGATCGCGC GTCTCGGTGT	1620

TTCCSCGSCC GGCATGACAA ACATGCAATG TGCATGACAA GTTACAGGTA TTAGGTCCAG	1680
GTTCACACAA GAGACAGGCA ACATGGGCTC ACCTTTTATG ACCGATCCGC ACCCGATGCC	1740
GGACATGSCG GGCCTTTTTC AGGTGCACGC CCAGACGGTG GAGGACGAGG CTCGCCCGAT	1800
GTGGGGCTCC GCGCAAAACA TTTCGGGTGC GGGCTGGAGT GGCATGCGCG AGGCGACCTC	1860
GCTAGACACC ATGGTTCAGA TGAATCAGGC GTTTCGCAAC ATCGTGAACA TGCTGCACGG	1920
GGTGCCTGAC GGGCTGCTTC GCGACGCCAA CAACTACGAG CAGCAAGAGC AGGCTTCCCA	1980
GCAGATCTTC AGCAGCTAAC GTCAGCGGCT GCAGCACAAT ACTTTTACAA GCGAAGGAGA	2040
ACAGGTTCSA TGACCATCAA CTATCAATTC GGGGATGTCC AGCTTCACGG CGCCATGATC	2100
CGGCTTCAGG CCGGTTTCTT GAGGCGCGAG CATCAGGCCA TCATTCGTGA TGTGTTGACC	2160
GCGASTGACT TTTGGGGCGG CCGCGGTTCC GCGGCTGCC AGGGGTTCTT TACCCAGTTC	2220
GGCGTAAGT TCCAGGTGAT CTACGAGCAG GCCAAGCGCC ACGGGCGAGAA GGTGCAGGCT	2280
GCCGCAACA ACATGCGGCA AACCGACAGC GCGCTCGGCT CCAGCTGGGC CTGACACCAG	2340
GCCAAGGCCA GGGACGTGCT GTACGASTGA AGTTCCCTCC GTGATCCTTC GGGTGGCACT	2400
CTAAGTGGTC AGTCTCGGGG TGTGCTGCTT TTCTCTCTTC GCGGTTCTT CGGTCTGTGT	2460
CACTGCTGCT GGGGCTCGGG TGAGGACCTC GAGGCGCAGG TAGGCGCGTC CTTCGATCCA	2520
TTCTCTGTCT TGTTCGCGCA GACGCTCTCC GACGAGGCGG ATGATCGAGG CGCGGTCCGG	2580
GAAGATGCCC ACSAGCTCGG TTGGGCTGCG TACCTCTCGG TTGAGGCGTT CTTGGGGGTT	2640
GTTCGACCGG ATTTGGCGCC AGATCTGCTT GGGGAAGGCG GTGAACGCCA GCAGGTGGGT	2700
GCGGCGCGTG TCGAGTCTCT CCGCCACCGC GGGGAGTTTG TCGGTCAGAG CGTCCAGTAC	2760
CCGATCAIAT TGGGCAACAA CTGATTCGCC GTGCGGCTGG TCGTAGATGG AGTCAGCCAG	2820
GGTGCSCACC CACGCGCAGG AGGCTTCGG GGTGCTGCC ATCAGATTGG CTGCTAGTGC	2880
GTTTCTGCAG CGCTCCGAGG CCGCTGCGGG CAGGCTGGCG CCGATCGCGG CCACGAGGCT	2940
GSCGTGGGCG TCGTGTGTA CCAGCGCGAC CCGGACAGG CCGCGGCGCA CCAGGTCCCG	3000
GAAGAAGGCC AGCCAGCTCG CCGCTCTCTC GCGGAGGTG ACCTGGATGC CCAGGATC	3058

(2) INFORMATION FOR SEQ ID NO:107:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 191 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single

123

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

```

Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
 1           5           10           15
Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
 20           25           30
Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
 35           40           45
Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
 50           55           60
Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
 65           70           75           80
Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
 85           90           95
Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
100           105           110
Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
115           120           125
Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
130           135           140
Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala
145           150           155           160
Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr
165           170           175
Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser
180           185           190
Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
195           200           205
Gln Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu
210           215           220
Gly Gly Leu Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn
225           230           235           240
Met Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val
245           250           255
Ser Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala
260           265           270

```

124

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala  
 275 280 285

Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly  
 290 295 300

Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val  
 305 310 315 320

Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg  
 325 330 335

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Gln Arg Gly Pro Gly  
 340 345 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly  
 355 360 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met  
 370 375 380

Pro His Ser Pro Ala Ala Gly  
 385 390

## (2) INFORMATION FOR SEQ ID NO:188:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1735 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (1A) SEQUENCE DESCRIPTION: SEQ ID NO:188:

GACGTCAGCA CCCCCCTGC AGGCTGAG AGTGCTCGT TTTGATCTGC GGTCAAGGTG 60

ACGTCCTCTG GCGTGTGCCC GCGGTGATG CAGACTCGAT GCGCTCTTTT AGTCCAACTA 120

ATTTCTTTGA AGTGCCTTCC AGGTATAGGA CTTCACGATT GGTAAATGTA GCGTTCACCC 180

CCTGTTGGGG TCGATTGGC CGGACCAATC GTCACCAACG CTTGGGCTGC GCGCCAGGCG 240

GGCGATCAGA TCGCTTGACT ACCAATCAAT CTTCAGCTCC CGGCGGATG CTCGGGCTAA 300

ATGAGGAGGA GCACGCTGT CTTCACCTGC GCAACGGAG ATGTTGGCGG CCGCGGCTGG 360

CGAAGTTCTT TCGGTGGGG CAACGCTGAA GGCTAGCAAT GCGCGGCGAG CCGTGGGCTG 420

GACTGGGGTG GTGCCCCCGG CTGCCAGCA GGTGTGCTG CTGCTTGCCA CACAATTCCG 480

TACGCATCCG GCGACGATC AGACGGCCAG GCGCAAGGCC GCGGTGATCC ATGAGCAGTT 540

TGTGACCAGG CTGCGCACCA CGGCTAGTTC ATATGCGGAC ACCGAGGCGG CCAACGCTGT 600

GGTCACCGCG TAGCTGACCT GACGCTATTC GACCGGAAGG ATTATCGAAG TGGTGGATT 660

```

CGGGGCGTTA CCACCGGAGA TCAACTCCGC GAGGATGTAC GCGGCGCCGG GTTCGGCCCTC      720
GCTGGTGGCC GCGCGCAAGA TGTGGGACAG CGTGGCGAGT GACCTGTTTT CGGCGCGCTC      780
GGCGTTTCAG TCGGTGGTCT GGGGTCTGAC GGTGGGGTCS TGGATAGGTT CGTCGGCGGG      840
TCTGATGGCG GCGGCGGCGT CGCGGTATGT GCGGTGGATG AGCGTCACCG CGGCGCGGGC      900
CCAGCTGACC GCGGCGCGAG TCGGGGTTGC TCGGCGCGCC TACGAGACAG CGTATAGGCT      960
GACGCTGCCC CGCGCGGTGA TCGCGGAGAA CGGTACGAA CTGATGACGC TGACCGCGAC      1020
CAACTTCTTG GCGCAAAACA CGCGCGCAT CAGGCGCAAT CAGGCGCAT ACAGCGCAT      1080
GTGGGCGCAA GACGCGGAGG CGATGTATGG CTACCGCGCC AGCGCGCGCA CGGCGCGCA      1140
GGCGTTGCTG CGTTGCGAGG AGCGCGCAT GATCAGCAC CGCGCGCGCG TCCTTGAGCA      1200
GGCGGTGCGG GTGAGGAGG CGGTGACAC CGCGCGCGCG AACCGTTGA TGAACAATGT      1260
GCGGCAAGCG CTGCAACAGC TCGCGCGCGC AGCGCGCGCG GTCGTACCTT CTTCGAAGCT      1320
GGGTGGGCTG TGGAGCGCGG TCTCGCGCGC TCTGTGCGCG CTCAGCAAGC TCAGTTGAT      1380
AGCGAACAAC CACATGTGGA TGATGCGCAC GGTGTGTGCG ATGACCAACA CTTTGCACTC      1440
GATGTTGAAG GCTTAGCTC CGCGCGCGCG TCAGCGCGTG GAAACCGCGG CGGAAAACCG      1500
GCTGTGCGCG ATGAGCTGCG TGGCAGCGCA GCTGGGTTG TCGGTGGGTT CTTCGGGTCT      1560
GGGCGCTGCG GTGGCGCGCA ACTTGGGTG GCGCGCGCTG GTCGGTTGCT TGTGGGTGCG      1620
GCGAGCATGG GCGCGCGCGA ACTAGCGCGT CAGCGCGCGG GCGCGCGCGG TCGCGGTGAC      1680
CAGCCTGACC AGCGCGCGCG AAACCGCGCG CGGACACATG CTGCGG      1725

```

(2) INFORMATION FOR SEQ ID NO:109:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

```

Val Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
 1                               5           10           15
Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Lys Met Trp
                20                   25           30
Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
        35                   40                   45

```

126

Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly  
 50 55 60  
 Leu Met Ala Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr  
 65 70 75 80  
 Ala Gly Gln Ala Gln Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala  
 85 90 95  
 Ala Tyr Glu Thr Ala Tyr Arg Leu Thr Val Pro Pro Pro Val Ile Ala  
 100 105 110  
 Glu Asn Arg Thr Glu Leu Met Thr Leu Thr Ala Thr Asn Leu Leu Gly  
 115 120 125  
 Gln Asn Thr Pro Ala Ile Glu Ala Asn Gln Ala Ala Tyr Ser Gln Met  
 130 135 140  
 Trp Gly Gln Asp Ala Glu Ala Met Tyr Gly Tyr Ala Ala Thr Ala Ala  
 145 150 155 160  
 Thr Ala Thr Glu Ala Leu Leu Pro Phe Glu Asp Ala Pro Leu Ile Thr  
 165 170 175  
 Asn Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Glu Ala Ile  
 180 185 190  
 Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu  
 195 200 205  
 Gln Gln Leu Ala Gln Pro Ala Gln Gly Val Val Pro Ser Ser Lys Leu  
 210 215 220  
 Gly Gly Leu Trp Thr Ala Val Ser Pro His Leu Ser Pro Leu Ser Asn  
 225 230 235 240  
 Val Ser Ser Ile Ala Asn Asn His Met Ser Met Met Gly Thr Gly Val  
 245 250 255  
 Ser Met Thr Asn Thr Leu His Ser Met Leu Lys Gly Leu Ala Pro Ala  
 260 265 270  
 Ala Ala Gln Ala Val Glu Thr Ala Ala Glu Asn Gly Val Trp Ala Met  
 275 280 285  
 Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu  
 290 295 300  
 Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser  
 305 310 315 320  
 Leu Ser Val Pro Pro Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro  
 325 330 335

127

Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Gln Thr  
 340 345 350

Ala Pro Gly His Met Leu Gly  
 355

## (2) INFORMATION FOR SEQ ID NO:110:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3027 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```

AGTTCACTCG AGAATGATAC TGACGGGCTG TATCCACGAT GGCTGAGACA ACCGAACCAC      60
CGTCGGACCG GGGACATCG CAAGTCGACG CGATGGGCTT GGCTCCCGAA GCGGAAGCCG      120
CCGAAGCCGA AGCGCTGCGC GCGCGGCGCG GGGCGGCTGC CGGTGCGCGC CGTTGAAGC      180
GTGAGGCGCT GCGATGCGC CCAGCCGAGG ACGAGAACCT CCCCAGGAT ATGCAGACTG      240
GGAAGACGCT GAAGACTATG ACGACTATGA CGACTATGAG GCGCGAGACC AGGAGGCGCG      300
ACGTCGCGCA TCTGCGCAC GCGCGTTGCG GTTCGCGTTA CCAAGACTGT CCACGATTGC      360
CATGCGGCTT GCAGTCCTCA TCATCTGCGG CTTCACCGGG CTCAGCGGAT ACATTGTGTG      420
GCAACACCAT GAGGCGACCG AACGCGAGCA GCGCGCGCGC GCGTTCGCGC CCGGAGCCAA      480
GCAAGGTCCT ATCAACATGA CTTGCGTGGG CTTCAACAAG GCGAAGAAG AGTTCGCGCG      540
TGTGATCGAC AGTTCACCG GCGAATTCAG GATGACTTC CAGCAGCGCG CAGCCGATTT      600
CAGCAGGCTT GTCAACAGT CCAAGCTGCT CACCGAAGGC AGGCTGAACG CGACAGCGCT      660
CGAATCCATG AACGAGCATT GCGCGCTGCT GCTGCTGCGG GCGACTTCAC GGCTCACCAG      720
TTCCGCTGGG GCGAAGACCG AACCAGCTGC GTGCGGCGTC AAAGTGACCG TGACCGAAGA      780
GGGGGACAG TACAAGATGT CGAAGCTTGA GTTCGTACCG TGACCGATGA CGTACCGCAC      840
GTCAACACCG AAACCACTGA CCGCAGCGAA GTGCTGAGA TCGACTCAGC CGCAGGCGAA      900
GCGGCTGATT CCGGACCGA GGCATTTGAC ACCGACTCTG CACCGAATC TACCGCGCAG      960
AAGGCTCAGC GCGACCGTGA CTTGTGCGCA ATGCGAGTTA CTTGAAACC GTTCCGCTG      1020
ATTCTCATCC TGCTCATGTT GATCTCTGCG GCGCGACCG GATGCGTATA CTTGAGCAA      1080
TACGACCGCA TCAGCAGACG GACTCGGCGC GCGCGGCTGC TCGCTGCGC GCGGCTCTG      1140
ACGGACCAAT CGCGCTCTTG TGTATTGACC GACAGCTCG ACCAAGACTT CGTACCGCC      1200

```

AGGTGGCACC TGGCCGCGCA TTTCCTGTCC TATACGACCA GTTCACGCAG CAGATCGTGG	1260
CTCCGCGCGC CAAACAGAAG TCACTGAAAA CCACCGCCAA GGTGGTGCGC GCGGCGGTGT	1320
CGGAGCTACA TCCGGAATCG GCGGTGTTT TGGTTTTTGT CGACCAGAGC ACTACCGTA	1380
AGGACAGCCC CAATCCGTGG ATGGCGGCCA GCAGGTGAT GGTGACCTTA GCCAAGGTGG	1440
ACGGCAATGG GCTGATCACC AAGTTCACCC CGGTTTAGGT TGGCGTAGGC GGTGCGCAAG	1500
TCTGACGGGG GCGCGGGTGG CTGCTCGTGC GAGATACCGG CGTTCTCCG GACAAACAGC	1560
GCCCGACCTC AAACAGATCT CGCGCGCTGT CTAATCGGCC GGGTTAATTA AGATTAGTTG	1620
CCACTGTATT TACCTGATGT TCAGATTGTT CAGCTGGATT TAGCTTCGGG GCAGGCGCGC	1680
TGGTGCATTT TGCATCTGGG GTTGTGACTA CTTGAGAGAA TTTGACCTGT TGGCGACCTT	1740
GTTCGCTGTC CATCATGGGT GCTAGTTATG GCGAGCGGA AGGATTATCG AAGTGGTGG	1800
CTTCGGGGCG TTACCACCGG AGATCAACTC CGCGAGGATG TACGCCCGCC CGGTTTCGGC	1860
CTGGTGGTG GCGCGCGCGA AGATGTGGGA CAGCTGCGG AGTGACCTGT TTTCGGCGCG	1920
GTGGCGTTT CAGTGGGTGG TGTGGGTCT GACGACGGA TGGTGATAG GTTGGTGGC	1980
GGTCTGATG GTGGCGCGG CTTGCGCGTA TGTGGGTGG ATGAGCTCA CCGCGCGCA	2040
GCGCGAGCTG ACCGCGCGCC AGGTCCGGGT TGCTGCGCG GCGTACGAGA CGGCGTATGG	2100
GCTGACGGTG GCGCGCGCG TGAATCGCGA GAACGTGCT GAACTGATGA TTCTGATAGC	2160
GACCAACCTC TTGGGCGAAA ACACCGCGCC GATCGCGCTC AACGAGCGCC AATACGGGGA	2220
GATGTGGGCG CAAGACCGCG CGCGATGTT TGGCTACCG GCCACGGCG CGACGCGGAC	2280
CGAGGCTTTC CTGCGTTTC AGGACGCGCC ACTGATCACC AACCGCGCG GGTCTCTTGA	2340
GCAGGCTTTC GCGGTGAGG AGGCGATCGA CACCGCGCG GCGAACGAGT TGAATGAACAA	2400
TGTGCGCGAA GCGCTGCAAC AACTGCGCG GCGACGAAA AGCATCTGGC CTTTCGACCA	2460
ACTGAGTGAA CTCTGGAAAG CCATCTCGCC GCATCTGTTC CCGCTGAGCA ACATCGTGTG	2520
GATGCTCAAC AACCACTGT GATGACCAA CTCGGTGTG TCGATGGCGA GCACCTTGCA	2580
CTCAATGTTG AAGGGCTTTG CTCCGCGCG GGTTCAGCG GTGGAAACCG CGGCGCAAAA	2640
CGGGGTCCAG GCGATGAGT CCGTGGCGAG CCAGCTGGGT TCGTCCGTGG GTTCTTGGG	2700
TCTGCGCGCT GGGGTGCGCG CCAACTGGG TCGCGCGCG TCGGTGGTT CTTTGTGGT	2760
GCGCGAGCGT TGGCGCGCG CCAACGAGC GGTACCGCG GCGCGCGCG CCGTCCGCT	2820